

DEZA DIREKTION FÜR ENTWICKLUNG UND ZUSAMMENARBEIT
DDC DIRECTION DU DÉVELOPPEMENT ET DE LA COOPÉRATION
DSC DIREZIONE DELLO SVILUPPO E DELLA COOPERAZIONE
SDC **SWISS AGENCY FOR DEVELOPMENT AND COOPERATION**
COSUDE AGENCIA SUIZA PARA EL DESARROLLO Y LA COOPERACIÓN



Eidg. Departement für auswärtige Angelegenheiten
Département fédéral des affaires étrangères
Dipartimento federale degli affari esteri
Swiss Ministry of Foreign Affairs
Ministerio Suizo de Asuntos Exteriores

Emploment and Income Division

phone +41 31 322 10 65
fax +41 31 323 08 49
Email: e-i@deza.admin.ch
www.sdc.admin.ch/themes

Current State of the Art in Impact Assessment: With A Special View on Small Enterprise Development

Author: Dr Brigitte Späth, Consultant
Kreuzlingen, Switzerland
August 2004

0. Introduction	3
1. Some General Considerations on Impact Assessment	3
1.1 Some matters of clarification	3
1.1.1 <i>Who wants an assessment and what for?</i>	3
1.1.2 <i>What is impact assessment?</i>	4
1.1.3 <i>What level of impact?</i>	4
1.1.4 <i>The impact hypothesis and attribution gap</i>	5
1.1.5 <i>What type and unit of impact?</i>	6
1.2. How can impact be assessed?.....	7
1.2.1 <i>A quantitative or “scientific” statistical method</i>	7
1.2.2 <i>A qualitative method</i>	8
1.2.3 <i>A participatory learning and action method</i>	8
1.2.4 <i>A proving - improving dichotomy</i>	9
2. Assessing Impact in Small Enterprise Development - SED.....	9
2.1 The BDS-Approach to SED	10
2.2 A Performance Measurement Framework - PMF.....	11
2.3. Assessing Poverty Impact of Small Enterprise Initiatives.....	11
2.3.1 <i>The need for defining a poverty line</i>	12
2.3.2 <i>Four Poverty indicators</i>	12
2.3.3 <i>Three impact chains for reducing poverty through BDS intervention</i>	13
2.3.4 <i>Three methods for assessing impact</i>	14
2.3.5 <i>Data source</i>	14
2.3.6 <i>Four recommendations</i>	15
2.3.7 <i>Critical appreciation of Oldsman’s paper</i>	15
2.4 Enterprise Development Impact Assessment Information Services (EDIAIS).....	17
2.5 Lessons from micro-finance?	18
2.5.1 <i>Assessing the Impact of Micro-enterprise Services (AIMS) research project</i>	18
2.5.2 <i>The Consultative Group to Assist the Poorest -CGAP</i>	21
2.5.3 <i>Lessons learnt</i>	22
3. Alternative ways to look at impact (assessment).....	23
3.1 Orientation towards Results: GTZ’s New Impact Monitoring.....	23
3.2 Six Steps to impact monitoring	24
3.2.1 <i>Identification of system borders</i>	24
3.2.2 <i>Agreement upon objectives of the monitoring</i>	25
3.2.3 <i>Identification of the areas of impact & formulation of impact hypotheses</i>	25
3.2.4 <i>Development and selection of indicators</i>	25
3.2.5 <i>Selection of assessment method & implementation</i>	26
3.2.6 <i>Capitalization of information and feed-back</i>	26
3.2.7 <i>Summing up</i>	26
3.3 A participatory learning impact assessment method - Method for Impact Assessment of Poverty Alleviation Projects (MAPP)	27
4. Concluding remarks.....	29
Annex 1: Websites of Interest on Impact Assessment.....	33
Annex 2: Overarching development goals, MDGs, CDF, PRSP, PSIA etc.....	35
Annex 3: Checklist for preparing an impact assessment.....	37
RELEVANT PAPERS AND ARTICLES	38

0. Introduction

In recent years impact assessment has become an increasingly important aspect of development activities. This is spurred by questions like: Do we do the right things? Do we do them right? The field of small enterprise development (SED) is no exception.

Although, there is an overall concern for impact assessment in development cooperation, each “compartment” or “discipline” seems to need its own specific approach and methodology to assess impact. This applies also for the wider field of SED. For example in the field of micro-finance, including financial services for small enterprises, the issue has been dealt with since some time (cf. 2.5 below). Currently, attempts are under way to come up with something specific for the field of non-financial services to small enterprises, and in particular to the market development approach for Business Development Services (BDS) .

While there are certain peculiarities in the different fields of development assistance which require specific attention, the wider task of impact assessment has, nevertheless, a lot of communalities. This paper will first address the common issues of impact assessment. Secondly, it will look at the approaches and methods developed and discussed particularly for SED and related development interventions. And finally it will introduce some other concepts for looking and assessing impact.

1. Some General Considerations on Impact Assessment

Impact assessment (IA) is as “...much an art as a science...” (according to Little, 1997, p. 2, cited in Hulme 2000, p. 79).

1.1 Some matters of clarification

A starting point of any impact assessment should actually be to reach a common understanding about some crucial questions and to make informed choices with regard to impact assessment.

1.1.1 Who wants an assessment and what for?¹

Different actors, such as donors, executing agencies, facilitating organisations in the partner countries, service providers, target groups etc. have different interests.² The donor agencies’ interest in “*proving*” predominated for a long time. There is increasing recognition that impact assessment does not only provide information for improving the accountability and advocacy position of development cooperation but also that it is useful to practitioners and management. This has led to a broadening of the definition and interpretation of impact assessment. Subsequently, impact assessment moved gradually away from narrow, donor focused impact assessment events, towards more practitioner-focused processes.³ Hence, there is scope for using impact assessment processes to tell us much more. The range of objectives can be:

- accountability: external and internal proving;

¹ This sub-chapter draws much from the www.imp-act.org website on context and objectives. Imp-Act stands for Impact Assessment of Microfinance: an Action Research Program, cf. also Annex 2.

² This goes in line with a debate on who should carry out what type of performance measurement and impact assessment and who should pay for these activities, cf. Miehlebradt/McVay 2002, p. 105.

³ „Monitoring and evaluation (M&E) of development activities provides government officials, development managers, and civil society with better means for learning from past experience, improving service delivery, planning and allocating resources, and demonstrating results as part of accountability to key stakeholders. Within the development community there is a strong focus on *results* — this helps explain the growing interest in M&E. Yet there is often confusion about what M&E entails” (World Bank 2002, p. 5).

- support advocacy & public relation;
- improving & learning: management, approach, instruments, methods;
- reinforcing institutional capabilities including self-analysis skills;
- strengthening the relationship between stakeholders; etc.

These objectives are not mutually exclusive, and many organisations will want to achieve a balance between a number of objectives in their impact assessment.

1.1.2 *What is impact assessment?*

There is, however, some confusion about what is meant by impact assessment and its related activities. Impact assessment studies have become increasingly popular with donor agencies and, in consequence, have become an significant activity for recipient agencies. In part this reflects a cosmetic change, with the term impact assessment simply being substituted for evaluation (cf. Hulme, 2000 p. 80). According to a handbook for practitioners, published by the World Bank, the following explanations are given: “A comprehensive evaluation is defined in the literature as an evaluation that includes monitoring, process evaluation, cost-benefit evaluation, and impact evaluation. Yet each of these components is distinctly different. (...) impact evaluation is intended to determine more broadly whether the program had the desired effects on individuals, households, and institutions and whether those effects are attributable to the program intervention. Impact evaluation can also explore unintended consequences, whether positive or negative, on beneficiaries” (Baker, 2000, p. 1).

The foreword of the recently published “Glossary of Key Terms in Evaluation and Result based Management” of OECD’s Development Assistance Committee (DAC) it is stated: “Evaluation is a field where development partners – often with widely differing linguistic backgrounds – work together and need to use a common vocabulary. Over the years, however, definitions evolved in such a way that they bristled with *‘faux amis’*, ambivalence and ambiguity. It had become urgent to clarify and refine the language employed and to give it a harmonious, common basis” (OECD/DAC 2002, p.4). The OECD/DAC terminology is not yet universally applied; there can be considerable differences even with one country.⁴ Such variations in terminology are closely linked to differences in the planning and project management approach of each organization but perhaps also to different organizational cultures.

1.1.3 *What level of impact?*

There are different (planning) levels where impact can occur. In the OECD/DAC glossary a distinction is made between:

- **impact** defined as “positive and negative, primary and secondary long-term effects produced by a development intervention, directly or indirectly, intended or unintended”;
- **outcome** defined as “the likely or achieved short-term and medium-term effects of an intervention’s outputs”;
- **output** defined as “the products, capital goods and services which result from a development intervention; may also include changes resulting from the intervention which are relevant to the achievement of outcomes” (cf. OECD/DAC 2002).

⁴ This is an observation which was also made in the frame of a cross-sectional analysis “Monitoring Instruments in German Development Co-operation” (BMZ, 2003) cf. www.dac-evaluations-cad.org/dac.abstracts_e.htm

The time perspective seems to be a crucial distinction between impact and outcome. All three - output, outcome or impact - are considered to be **results** of a development intervention. In SDC's own "glossary of the 27 most relevant terms related to evaluation and controlling" the definitions match closely with the one of OECD/DAC.

Figure 1: Levels of results according to SDC



Until recently most donors and their executing agencies used the logical framework (logframe or ZOPP) as a major planning and management instrument. Consequently, the terminology and planning levels of the logical framework became also a major reference in all attempts for measuring or monitoring performance and impact. This is illustrated by figure 1. SDC distinguishes between three result levels: (i) impact, (ii) outcome and (iii) outputs each of which is related to a different planning level. Accordingly, SDC locates impact

at the level of the goal or development objective, whereas outcome is positioned at the level of the objective of the program or project and outputs at a level lower.

1.1.4 The impact hypothesis and attribution gap

Behind all development initiatives is the assumption that intervention will change human behaviours and practices in ways that lead to the achievement of desired outcomes. In brief, all development programs are based on a working hypothesis, which is usually expressed in a more or less complex "cause" and "effect" relation. This causal relation between project input and output, followed by intermediate outcome and final outcomes is not only considered as a base for planning (logical framework / ZOPP) of a project intervention, but also for its monitoring and final evaluation. As a consequence, the development of a "causal development model" became a centre of concern for the Working Group for Impact Measurement and Performance (WIMP) of the Donor Committee. Kris Hallberg, a World Bank representative to the Donor Committee, recently stressed again the importance of the logical framework and its causal relation with a view to impact assessment.⁵

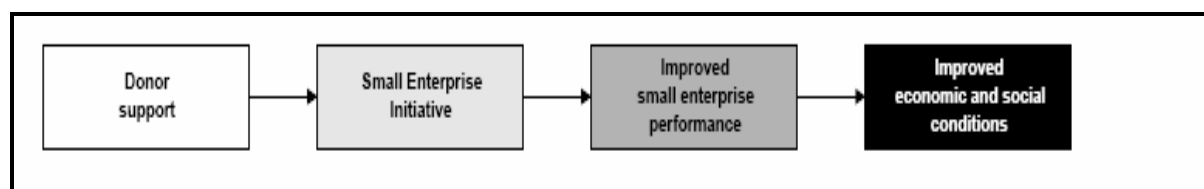
Based on the research findings on the Performance Measurement Framework (PMF) of BDS it was proposed to elaborate on a causal development model.⁶ This was done so by Hallberg and Oldsman who elaborated a "Framework for Evaluating the impact of Small Enterprise Initiatives" (2002). This methodological paper outlines a general approach to impact evaluation applicable to different types of SED interventions such as BDS market development, financial services and policy advice. Particular emphasis is put on the use of logical planning models to structure the chain of causality between program inputs, outputs, and outcomes. A

⁵ This was done so at the occasion of the BDS-Seminar in Turin, Sept. 2003, cf. The BDS Seminar 2003, Testing the Guiding Principles, 4th Annual Seminar 8-12 September 2003, Turin Italy, Seminar Report by Marieke de Ruijter de Wildt, ILO-TC, 2003 p. 25.

⁶ Cf. chapter 2.2 of this paper and McVay, Miehlbradt, and Canedo (2001 p. 4f.).

first crude causal model for assessing impact of SED-intervention was developed by Hallberg and Oldsman 2002.

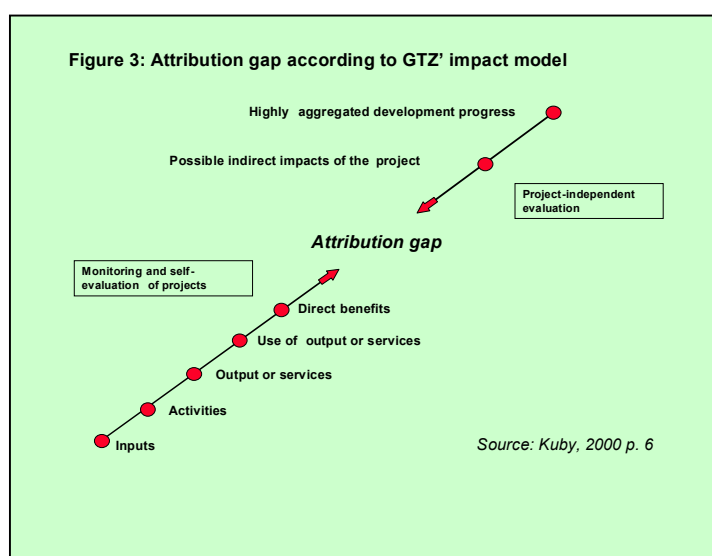
Figure 2: Causal Model for SED Impact



Source: Hallberg and Oldsman, 2002, p. 5 .

However, this causal model is not without problems. This is elaborated in more detail under 2.3.7. The fact that many (confounding) 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. This constitutes a fundamental methodological problem of impact analysis.

The existence of an “attribution gap” was first addressed by GTZ (cf. Kuby 2000). The gap divides the “impact chain” into two parts: one in close proximity to the project, but far away from development results, and another one up in the realms of overall development progress. For these reasons, other donors, particularly GTZ, have lately abandon the rigidity of the logical framework and have turned to a more open and flexible approach (cf. chapter 3.1).



1.1.5 What type and unit of impact?

Moreover, there are different types of impact. Conventionally, economic indicators have dominated SED impact assessment so far. With evaluators particularly keen to measure changes in income despite the enormous problems this presents.⁷ Other popular variables have been levels and patterns of expenditure, consumption and assets. Social indicators (e.g. educational status, access to health services, nutritional levels etc.) became popular in the early 1980s. With the Rio Summit (1992) the notion of ‘sustainable development’ was first coined, linking economic, social and environmental development aspects. In addition, the array of variables has recently been extended into the socio-political arena in an attempt to assess whether a development intervention promotes e.g. empowerment or gender equality.⁸

⁷ This statement was originally made for micro-finance (cf. Hulme 2000 p. 83) but it is also applies for non-financial assistance to small enterprises.

⁸ This has led to the measurement of individual control over resources, involvement in household and community decision making, levels of participation in community networks and social network etc. cf. Hulme, 2000, p. 83f.

These extension to the types of impact assessed sheds light on developmental impact at a time when the goals of development have also been expanded.

An ever increasing list of international conventions, action plans, goals etc. were declared over the last decades. One of the most prominent and recent commitments refer to the UN-Millennium Development Goals (MDGs) of which goal number 1 to “eradicate poverty and hunger” is the most well-known (cf. figure 4). Emphasis on poverty eradication has led to an increased focus on poverty impacts of development programs.⁹

<p>Figure 4: MDG 1 Goal 1: Eradicate extreme poverty and hunger Target 1: Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day Source: (www.undp.org/mdg).</p>	<p>Indicators</p> <ol style="list-style-type: none"> 1. Proportion of population below \$1 per day (PPP-values) 2. Poverty gap ratio [incidence x depth of poverty]
--	--

An almost infinite array of variables can be identified to assess impacts of different units. Likewise, the units or levels of assessment are selected according to the impact hypothesis of a development intervention as well as to the type of impact. Typical units of assessment are at the level of individuals, enterprises, households, communities, institution, etc. (Hulme 2000, p. 82f.).

1.2. How can impact be assessed?

In principle there are only three major approaches to impact assessment, each having its own merits and weaknesses.¹⁰

1.2.1 A quantitative or “scientific” statistical method

The quantitative method seeks to ensure that effects can be attributed to causes through experimentation. This is the domain of econometricians and statisticians and the tool applied is usually a survey. Assessing the impact of a development intervention hinges on asking a fundamental question: *What would the situation have been if the intervention had not taken place?* Although one obviously cannot observe such a situation, it is possible to approximate it by constructing an appropriate counterfactual. Counterfactuals are simulated by comparing program participants (treatment group) with a control or comparison group. The choice of method to identify the group of non-participants determines the evaluation design. One can broadly distinguish between three categories of evaluation design:

- (i) *experimental or randomized evaluation design:* This design involves gathering a set of individuals (or other units of analysis) equally eligible and willing to participate in the program and randomly dividing them into two groups.
- (ii) *quasi-experimental evaluation design:* The evaluation design consist of constructing a comparison group using matching or reflexive comparisons.
- (iii) *non-experimental evaluation design:* This evaluation design can be used when the other two methods cannot be applied. In such situations, program participants can be compared to non-participants using statistical methods to account for differences between the two groups.

⁹ For more information on MDGs, CDF, PRSP, PSIA etc, cf. Annex 2.

¹⁰ Cf. Hulme 2000, p. 84 f, Baker, 2000, p. 4 ff.; also www.worldbank.org/poverty/impact/

The focus is to get “robust” data that informs on *before or after*, or *with and without* the intervention. It is, however, quite tricky to net out the program impact from the counterfactual conditions that can be affected by history, selection bias, and contamination (cf. Backer 2000 p. 2). Moreover, the information provided do give no indication on what is behind the “numbers” while the amount of resources required for such a “scientific survey” in terms of specialist advice, surveyors, money and time for gathering and processing the data is quite high (cf. Hulme, 2000, p. 88, World Bank 2000, p. 12).

1.2.2 *A qualitative method*

The qualitative method is rooted in “humanities tradition” such as sociology, anthropology etc.¹¹ Its main features are an inductive approach, a focus on key informants, recording by notes or image, and the data analyst is usually directly involved in data collection. It includes such tools as interviews, participants observations, case studies, focus group discussions, etc. Among the methodologies used in qualitative impact assessments are the techniques developed for rapid rural assessment, which rely on participants’ knowledge of the conditions surrounding the project or program being evaluated (Backer 2000, p. 8). Although, qualitative data can be quantified, this tradition does not try to “prove” impact within statistically definable limits of probability. Rather, it seeks to provide an interpretation of the processes involved in intervention and of the impacts that have a high level of plausibility (cf. Hulme 2000, p. 86). The validity of an evaluation or assessment established with such a method can be judged on the basis of the logical consistency of the arguments and materials presented; the strength and quality of the evidence provided; the degree of triangulation used to crosscheck evidence; and the quality of the methodology. However, such studies can usually not demonstrate a causal link as they are not able to generate a “without program” control group. Instead, causality is inferred from the information about the causal chain collected from intended beneficiaries and key informants, and by comparison with data from secondary sources about changes in “out-of-program” areas. The ability to gather qualitative information about processes, behaviours as well as the diversity of perception by high to medium skilled practitioners is high, whereas other resources (costs, time) required range from high to low (cf. Hulme 2000, p. 88, Word Bank 2000, p.14f.) .

1.2.3 *A participatory learning and action method*

In this method the stakeholders are involved in all stages of the evaluation or assessment, such as determining the objectives of the study, identifying and selecting indicators to be used, and participating in data collection and analysis.¹² It was developed out of a critique of the “external” and “scientific” evaluation which according to the proponents’ view fails to grasp the complexity, diversity and contingency of winning a livelihood. According to this line of argument the scientific methods “...reduces causality to simply unidirectional chains, rather than complex webs; it measures the irrelevant or pretends to measure the immeasurable; and it empowers professionals, policy-makers and elites, thus reinforcing the *status quo* ...” (Hulme 2000, p. 87). Moreover, it is recognised that there are many realities and before any analysis or action is taken the individuals concerned must ask themselves “whose

¹¹ cf. Hulme 2000, p. 85 f, Baker, 2000, p. 7ff.

¹² Cf. www.enterprise-impact.org.uk and particularly the documents of Linda Mayoux but also the papers presented at the conference NEW DIRECTIONS IN IMPACT ASSESSMENT FOR DEVELOPMENT: METHODS AND PRACTICE that took place in November 2003 in Manchester.

reality counts?”. The tools developed so far rely on participatory appraisals, action learning methods, etc. and it is assumed that the beneficiaries are empowered through the research process itself. Even though, the art of participatory impact assessment is still in its infancy, a number of comparative studies have shown that, when well conducted, participatory methods can be more reliable than conventional methods.

Although, there is an extensive literature on quantitative versus qualitative methods, and since recently also on participatory methods, in impact evaluation, there is also a growing acceptance of the need for integrating the different approaches.

1.2.4 A proving - improving dichotomy

Historically impact assessment was done to meet donors needs to justify funding. This is gradually changing in favour of a more learning and improving. There was a felt need to more clearly define the range of objectives to be fulfilled and questions to be answered through impact assessment, and the range of possibilities of approaches and tools available to meet these objectives.

Figure 5: Proving – Improving Opposites of Impact Assessment

	←-----→	
	Proving Impacts	Improving Practice
Primary Goal	Measuring as accurately as possible the impact of an intervention	Understanding the processes of intervention and their impacts so as to improve those processes
Main Audience	Academics and researchers Policymakers Evaluation departments Program Managers	Program managers Donor field staff NGO personnel Intended beneficiaries
Associated Factors	Objectivity Theory External Top down Generalisation Academic research Long timescales Degree of confidence	Subjectivity Practice Internal Bottom up Contextualisation Market research Short timescale Level of plausibility

Source: Hulme, D.: Impact Assessment Methodologies for Micro-finance: Theory, Experience and Better Practice, in: World Development Vol. 28, 2000, No. 1, p. 80

The objectives of impact assessment greatly influence the information needs, the approach and methodology applied. This is illustrated with a proving improving dichotomy of the primary goals of impact assessment above (cf. figure 5). The main audience of each type of assessment is listed as well as a set of factors associated with the extreme ‘pole’ positions of this continuum.

2. Assessing Impact in Small Enterprise Development - SED

Today, development cooperation in SED concentrates on establishing and improving structures – government, administrative, institutional, organisational, market, legal etc. – thereby facilitating the development of micro, small and medium scale enterprises. SED-practitioners argue quite often that measuring impact at the level of an overall development goal is so difficult because SED-interventions are not any more directly made at the level of the target group. However, this holds also for most other technical fields of assistance in official development assistance (ODA). Most development interventions are geared towards structural changes that will only in the medium and long run improve conditions in a sustainable way.

The OECD/DAC glossary (2002) clearly points to the relevance of interventions at the institutional level by separately defining “*institutional development impact*” as “The extent to which an intervention improves or weakens the ability of a country or region to make more efficient, equitable, and sustainable use of its human, financial, and natural resources, for example through: (a) better definition, stability, transparency, enforceability and predictability of institutional arrangements and/or (b) better alignment of the mission and capacity of an organization with its mandate, which derives from these institutional arrangements” (ibid). Exceptions are by definition to be found in the field of humanitarian aid and relief work, which have entirely different mandates and time frames. Other exceptions are usually programs for specific target groups and/or specific situations such as refugees, demobilisation of soldiers, youth, women, people in conflict regions as well as pilot activities engaged in developing and testing new instruments.

2.1 The BDS-Approach to SED

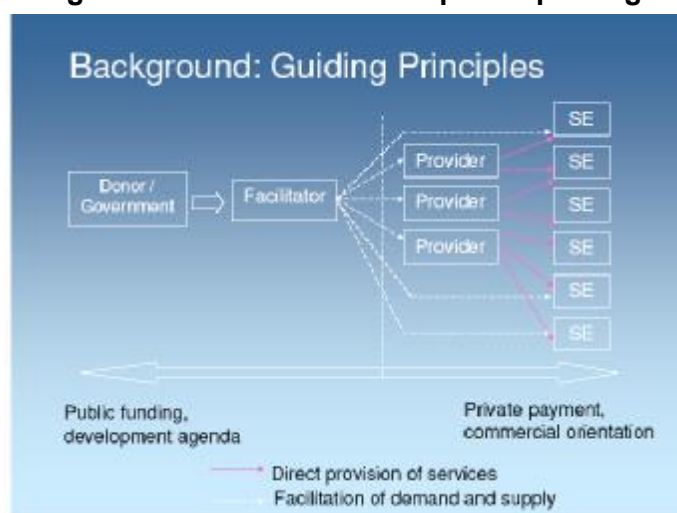
The necessity for assessing impact was perhaps spurred by the International Donor Committee for Small Enterprise Development and the promulgation of Business Development Services (BDS) as a new paradigm in the late 1990s. There is a need to provide evidence that the new paradigm is not only superior to other approaches, being it in terms of outreach, efficiency, sustainability but also in terms of contributing to higher development policy goals such as poverty reduction, employment and income generation, economic growth. In respect to small enterprise development based on the approach of Business Development Services (BDS) the Donor Committee has explicitly stated the following overarching goals:

„The ultimate objective of donor initiatives in Business Development Services (BDS) is to improve small enterprise (SE) performance in developing countries, as a means to achieve higher economic growth and employment, reduce poverty, and meet social objectives.”¹³

Usually, donor interventions in SED are directed towards national and local partner institutions which in turn may work indirectly or directly with the target group or clients. This mechanism of donor intervention has become more obvious and, perhaps, also more complex with the introduction of the BDS market development approach in SED, which is illustrated in figure 6.

The overall logic is that strengthening demand for and supply of services leads to a vibrant, competitive BDS market. BDS market development leads to large numbers of SEs using and applying business services to improve

Figure 6: BDS market development paradigm¹⁴



¹³ BDS for small enterprises. Guidelines for Donor Intervention, Donor Committee for SED 2000, p. 4

¹⁴ Source: Opening Presentation of Jim Tanburn, ILO at the SDC Workshop: “Opportunities and limits of the BDS approach in rural areas. Some Starting Points” Gerzensee, 8 January 2004.

their business practices. These changes in business practices contribute to SE growth and profitability, which in turn contributes to increased employment and income” (Miehlbradt and McVay 2002, p. 91). The donor’s immediate interaction is usually limited to a facilitator. As a consequence, the so-called impact chain - that is between a distinct project activity and a visible result - gets longer, and longer. And, it becomes more difficult to attribute impact to a distinct donor intervention.

2.2 A Performance Measurement Framework - PMF

As donors and executing agencies get along with applying the market development approach to BDS, performance measurement became more and more relevant. In a collaborative effort a common performance measurement framework (PMF) was developed. It resulted in a preliminary framework, the PMF 2001.¹⁵ It presents a logic causal model that attempts to include BDS market development goals. There is consensus on that programs should typically measure performance in the following categories:¹⁶

- **Impact** on small enterprises and the wider economic/social environment;
- **Outreach**, meaning both the number of SEs reached (scale) and the effort to provide services to people not served by existing markets (access);
- **Sustainability** of business service delivery and supplier institutions; and
- **Cost-Effectiveness** of program activities.

However, there remains considerable debate over how to measure the performance of BDS programs, and even more on assessing their impact. Performance assessment looks primarily at program outputs and outcomes and asks how many people use the service and how they use it; whether they are satisfied with it; and if it is sustainable. Depending on the program, different levels of intermediate outcomes are also assessed, tending to focus on outreach (number of people served and who) and outputs such as services delivered.

To some extent the PMF has even tried to venture into the field of higher impact assessment. There is a lot of interest in understanding and measuring impact on small enterprises and the wider economic and social environment. In impact, one looks at how changes in business practices, caused by business services, help businesses improve, create wealth and income for owners, and contribute to broader development goals such as job creation, economic development, empowerment, environmental preservation and more. However, it was realized that understanding the impact of program on small enterprises and the wider environment requires expensive and rigorous studies that can only be performed infrequently. The PMF does currently not, however, consider the ultimate impact on businesses, individuals, and communities. For regular monitoring, experts are exploring the use of proxies, indicators that are easier to measure than final impact but are linked to final impact.

2.3. Assessing Poverty Impact of Small Enterprise Initiatives

The Working Group for Impact Measurement and Performance (WIMP) of the Donor Committee contracted Eric Oldsman with the task to develop a framework for measuring the poverty impacts of small enterprise initiatives. The paper (Oldsman 2003) describes various

¹⁵ The PMF is an ongoing global initiative to identify valid, practical, and useful performance indicators for business development service (BDS) programs that serve small enterprises (SEs). Cf. McVay, Miehlbradt, and Canedo 2001; Miehlbradt/McVay 2002 and 2003.

¹⁶ Source: BDS Turin 2002: Seminar Reader – Miehlbradt & McVay p. 89

econometric methods that could be used to determine the extent to which a particular program reduces poverty within a target population. It concentrates on BDS-programs, although Oldsman states, that the principles and methods for measuring poverty impact discussed in the paper are widely applicable (p. 3).

At the beginning Oldsman refers to the overarching interest of donors to enhance the performance of small enterprises as a means of accomplishing various social and economic objectives, including poverty reduction. At the same time he points to the fact that no definition of poverty was made so far in the BDS-documents.

2.3.1 *The need for defining a poverty line*

Oldsman strongly emphasizes the necessity for establishing a threshold level of income or consumption below which people are considered poor if poverty is to be measured (p. 5, p.9). He suggests to take the income level as the most common definition for poverty: "In the absence of a good alternative, donor agencies may want to use \$1 per day (PPP)¹⁷ as the relevant poverty line for evaluations of SE initiatives since this would be consistent with the Millennium Development Goals" (ibid p. 10, footnote 22). However, "(t)he choice of which poverty line(s) to use in an evaluation of a particular BDS-program should be based on the objectives of the program as well as social norms as to what constitutes minimum living conditions"(ibid p. 10).

2.3.2 *Four Poverty indicators*

Evaluators will need to select one or more indicator that could be used in determining whether an initiative reduced poverty among a particular population (ibid p.9). For this purpose four different econometric poverty indicators are introduced and briefly discussed (ibid p. 9):

1. *"poverty rate (headcount):* (...) It is simply the proportion of the population living in households with income per person below a specified poverty line. (...)
2. *poverty gap index:* The index provides a measure of the depth of poverty within a population in terms of the gap between income or consumption levels and the specified poverty line. (...)
3. *squared poverty gap index:* This measure takes the severity of poverty into account by placing a higher weight on individuals living in households that are furthest from the poverty line.
4. *Sen-Shorrocks-Thon Index:* The final measure attempts to combine the incidence, depth and severity of poverty into one indicator (...)." (Oldsman 2003, p. 9, cf. also Appendix A).

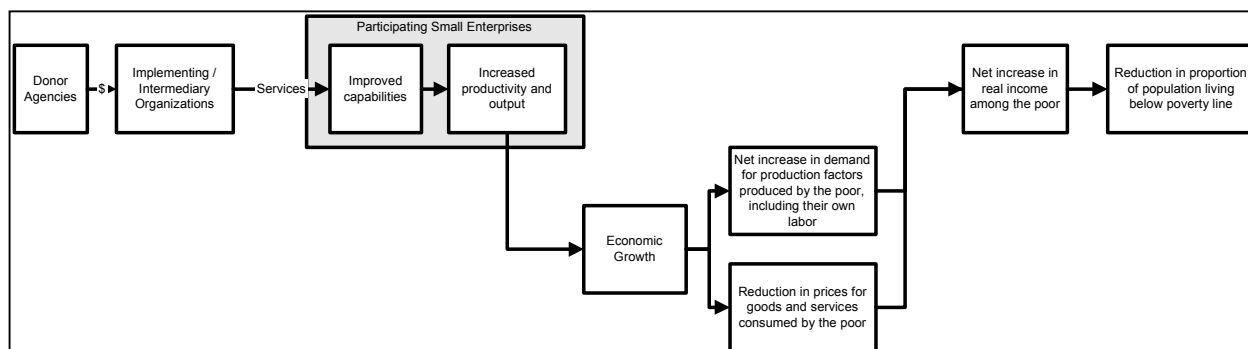
"As can be seen, a variety of indicators can be used to measure the incidence, depth and severity of poverty among different populations. (...) Multiple measures would provide more insight into the nature of impacts, providing information on whether programs reduced the proportion of the population living in poverty, reduced the overall poverty gap, and/or reduced inequality among the poor"(Oldsman 2003 p. 10).

¹⁷ The World Bank has introduced the a set of exchange rates that is called *purchasing power parity* (PPP) exchange rates which are designed to convert currencies in a way that preserves purchasing power, and which for comparison between India and the US, converts 454 rupees to around \$ 50 per person per month (Deaton, 2004 p. 17). Cf. also Pogge 2003).

2.3.3 Three impact chains for reducing poverty through BDS intervention

Based on the model developed earlier by Halberg/Oldsman (2002), the logical chain of causality between donor intervention and poverty reduction is developed further (cf. the example of such a chain below).

Figure 7. Poverty Impact through Economic Growth Generated by Participating Small Enterprises



Oldsman suggests that there are basically three different chains of impact for poverty reduction in BDS-programs (p. 5f). Poverty impact occurs:

1. when *the owners of participating small enterprise are poor*: Some BDS-programs target specific groups for intervention such as poor farmers or petty traders in a particular area. E.g. they help them to start or operate micro-enterprises. Because these programs work with business owners that are poor, improved business performance may yield sufficient income to lift these business owners (and unpaid family workers) out of poverty. These programs aim to reduce poverty within these specified groups.
2. when participating small enterprises *employ the poor*: In deed, most BDS-programs are not designed to address business owners who are themselves poor. Those that focus on the development of BDS-markets are not aimed at particular groups of poor people; rather, these programs are intended to reduce poverty within the general population. In these instances, the links between programs and poverty reduction is more complex, hinging to a great extent on induced labour demand within participating firms and the broader economy.
3. *through economic growth* generated by participating small enterprises: The improved performance of participating small enterprises may spur economic growth leading to higher real household income and a concomitant reduction in the poverty rate (cf. figure 7). SE initiatives benefit the poor to the extent that they result in pro-poor economic growth, i.e., growth that provides greater opportunities for the poor to generate higher real income. In this regard, numerous empirical studies have found that economic growth is associated with poverty reduction.¹⁸

¹⁸ Oldsman refers to Dollar/Kraay (2002) who published a paper for the World Bank with the title "Growth is good for the poor" in 2000. However, the scientific rigidity of the methodology applied was criticised by Lübker, Smith and Weeks 2002. Moreover, there is actually an ongoing discussion in what way the poor profit actually from overall economic growth, cf. e.g. Warr 2000. A most recent contribution to the debate is from Ranis (May 2004) on the interrelationship between human development and economic growth.

Cf. also Beck/ Demircuc-Kunt/Levine (2003) in another World Bank study on: The Impact of SMEs on Growth, Development, and Poverty: Cross-Country Evidence.

All three paths end with a reduction in the proportion of the population living below the poverty line. This is based on the assumption that, improved business performance reduces poverty to the extent that it raises the income of the poor. Regardless of the path, from an evaluation perspective, the ultimate question is whether donor-supported programs have actually closed the poverty gap and/or reduced the proportion of the population that lives in poverty. As such, evaluations need to be based on valid indicators of changes in poverty. According to Oldsman, indicators such as changes in output and productivity, employment levels and owner income of participating companies are of questionable validity as proxies for poverty impacts (p. 8).

2.3.4 *Three methods for assessing impact*

According to Oldsman (2003), the fundamental principle of impact assessment is the need to compare the observed situation with the intervention to *what would have been* had there been no intervention at all, i.e., the counterfactual. Valid comparisons imply that the net effect of interventions is isolated from all other extraneous or confounding factors that influence defined outcomes. Thus, the major challenge in impact assessments is to estimate the effect of programs after netting out extraneous factors that affect outcomes (ibid p.8). For this end, Oldsman suggests that the menu of options available to assess the impact of small enterprise initiatives on poverty is limited to the three basic methods:

1. "*Experiments with random assignment*: The gold standard in impact assessment is experimental design in which poverty rates in regions served by the small enterprise initiative are compared to similar regions that have not received assistance. The critical element of this design is the random assignment of regions to treatment and control groups" (ibid p. 12).
2. "*Quasi-experiments with constructed controls*: Quasi-experimental designs involve comparing the performance of poverty rates in regions served by the small enterprise initiative to similar regions that have not received assistance. Assignment to the two groups is non-random. Rather, a comparison group is constructed after the fact" (ibid p. 12).
3. "*Participant judgment and expert opinion*: This approach relies on people who are familiar with the intervention to make judgments concerning its impact" (ibid p. 12).

The different use, advantages and challenges of the three methods are discussed in more detail (ibid, p. 13-18). According to Oldsman, "the strength of causal inferences that can be drawn from the analysis depends on how well the particular approach used in assessing impacts deals with threats to validity" (p. 11).

2.3.5 *Data source*

In addition, attention is drawn to data sources. The type of impact assessment proposed by Oldsman requires accurate and reliable data: "At a minimum, it is necessary to obtain data on income and consumption for households before and after the SE intervention. Data for other control variables will also be needed depending on the precise model specification" (ibid p. 18f.). Furthermore, Oldsman points to the possibility of using existing data such as from household surveys. However, he also considers, that the available data may not be sufficient for the evaluation of a particular program given its time coverage, timing and content (ibid, p. 19).

2.3.6 *Four recommendations*

Finally, Oldsman makes four recommendations that are to be considered when the impact of BDS-programs on poverty is assessed:

1. Undertake evaluations of poverty impacts only if the scale of programs warrants it (ibid p. 19).
2. Define the poor and specify the causal mechanisms that are expected to lead to poverty reduction (ibid p. 20).
3. Select regions for programs in an explicit manner (ibid p. 20).
4. Assess the impact of programs on intermediate objectives (ibid p. 20).

2.3.7 *Critical appreciation of Oldsman's paper*

Oldsman introduces and discusses basically a quantitative or scientific approach to impact evaluation (cf. 2.1 above, also Sinning, 2003). He strongly advocates to make use of one of the two econometric standard methods: (i) the experimental or randomized evaluation design and (ii) the quasi-experimental evaluation design for impact assessment. The first one is generally considered to be the most robust of the evaluation methodologies, while the second one is for pragmatic reasons applied more often in evaluation.

There are a number of problems involved in applying these quantitative methodology, such as:

- the enormous demand for data required,
- the great need of resources, in terms of expert knowledge, staff, time, finance etc.
- the danger of a biased sample selection;
- fungibility – that the services or resources provided as a part of a development intervention, is used for a different purpose than the originally intended one by the beneficiary - ;
- misspecification of underlying causal relationships (cf. Hulme 2000, p. 84f.; Dunn 2002;).

In the following remarks three issues will be discussed in more detail: (a) the problem of causality; (b) the concept of poverty and (c) the quality of information.

The **causal model** and the attempt to quantify the effects at a highly aggregated level is not without problems. Jeanne Downing, convenor of WIMP put it in plain words:¹⁹ "...the further one goes to the right of the causal model, i.e. from intermediate impact objectives like BDS market development and improved firm performance to larger social impacts like poverty (reduction), (...) – the more abstract the measurement and the harder 'it becomes to measure impact with confidence'". Consequently, Oldsman warns of attempts to simplify causality relations:

"...while tempting to simply measure firm-level performance gains, this approach will not provide valid estimates of impact in terms of poverty reduction. With the exception of poor business owners who receive services under the auspices of some BDS-program, in general, the ultimate intended beneficiaries of programs – the poor – may be affected by programs through a variety of indirect channels" (Oldsman 2003, p. 10).

¹⁹ Minutes of the Meeting, Sept. 14 – 15 2003 in Copenhagen, communicated via e-mail Dec. 3, 2003

Thus the observed changes at the level of the poor cannot be reliably and proportionately traced back to any one intervention or program. Other factors could have affected the observed changes. 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. Hence, Oldsman himself states critically:

“... the ability of most programs to reduce poverty in a particular area is likely to be limited given the scale of the problem and the relative magnitude of the intervention (i.e., the effect size of programs is likely to be small). This means that it will be difficult to identify an impact even if it exists. Large datasets – which are likely to be costly to compile – will be needed to estimate poverty impacts.²⁰ Some programs will be able to justify the expenditure; others won't. Unfortunately, there is no inexpensive method to generate valid and reliable estimates of the impact of small enterprise programs on the poor” (Oldsman, 2003, p.10).

Instead, he suggests (ibid p. 20) that greater attention should be placed on conducting rigorous evaluation of programs in terms of their impact on firm-level performance using experimental or quasi-experimental designs. This impact is considered to be necessary in order to achieve the ultimate goal of poverty reduction.²¹

There is much debate about who constitute **‘the poor’**. In general, the measurement of (extreme) poverty has its pitfalls (cf. Pogge, NZZ 4./5. Jan. 2003). Besides one can question: “...why do we say that someone who is just below the poverty line is poor, and thus a candidate for transfers and the special attention of the World Bank, while someone who is just above it, whether by sixpence or six ananas, need no help and can be safely left to their own devices? Even if we could precisely set the poverty line, and even if we could precisely measure each person's income, neither of which conditions are close to being met, it makes no sense to treat such similar people differently.” (Deaton 2004, p. 8).

A concentration on drawing a “poverty line”²² according to the level of income, as is suggested by Eric Oldsman, focuses only on the economic dimension of poverty. In the words of Amartya Sen: “The point is not the irrelevance of economic variables such as personal incomes, but rather their severe inadequacy in measuring many of the casual influences on the quality of life and survival chances of people” (cit. by Khawari 2004, p. 5). According to Sen (1999) poverty is seen as deprivation of basic capabilities, rather than simply low income. The international debate has moved away from the relatively narrow notion of income-metric measures to a wider understanding of poverty that takes into account questions of assets, vulnerability, diversity and participation. The concept of poverty is increasingly acknowledged as a multi-dimensional problem that involves more than deprivation of income, but also lack of freedom, increased vulnerability, risk, and powerlessness (cf. Nadvi/Barientos 2004, p. 5f.).

²⁰ The need for large amounts of data is driven, in part, by the interaction between small effect sizes and large measurement errors. Oldsman points here to the so-called small-N problem, cf. Neubert 2004b p. 9.

²¹ According to the authors knowledge this is actually done in some SDC funded programs that are executed by Swisscontact. In Peru: *Sistmatización de la experiencia y medición de impacto de los centros de desarrollo empresarial del sistema Swisscontact*, by Fernando Villarán (SASE) and Julio Gutiérrez (SASE), December 2000; and the Russian Federation: *Impact Assessment of the Swisscontact Foundation Activity in Voronezh*, Prepared for: Swisscontact Voronezh office by: Reforma Company January 2004.

²² For extreme poverty line below 1 US \$/day.

Quantitative data is usually considered as “**cold**” **information**. Only when being interpreted, that is usually with either the experience of a person (expert knowledge) or with the assistance of additional qualitative information, do they unfold their information potential (cf. SDC, 1999 p.3). Moreover, such type of impact assessment is usually an once-in-time exercise, so there is little scope for observing changes over time. From the perspective of a more process-oriented analytical framework, it is, however, essential that programs are to continually learn from their experience and improve (rather than prove) their impact (cf. Hulme, 2000 p. 85).

2.4 Enterprise Development Impact Assessment Information Services (EDIAIS)

In order to do justice to other existing initiatives and experiences in impact assessment with a special view to SED the following information is briefly given. EDIAIS, managed on behalf of DFID, has in its centre a webpage in which various ‘core texts’²³, tools and instruments, particularly in line with DFID requirements, are made accessible (cf. Annex 1). In this collection of definitely helpful instruments and tools, it is, however, difficult to recognize a guiding principle or overarching approach to impact assessment.

End of 2003 EDIAIS held a international conference on “New Directions in Impact Assessment for Development: Methods and Practice”²⁴. However, the papers presented and discussed at this conference were not limited to SED. The plenary and working group sessions were divided in:

- *Strategic level*, which addressed more the overarching development policy goals such as assessing sustainable development integrated impact assessment or the Poverty and Social Impact Assessment (PSIA) designed by the World Bank;
- *Program level*, concentrated on assessing impact of programs and projects as part of the management. Among others a network analysis was presented as an alternative to the widely criticised but still dominating logframe method for planning and documentation.
- *Grassroots level*, focused on participatory learning and/or observation approaches, that involves reversing and changing mindsets. Strong points were made on the focus of proving impact rather than reliability of the conventional approaches of assessment. It was critically noted that participation is still frequently neglected by failing to engage participants either in designing questions or in analysing results (cf. Enterprise Impact News Issue 26, Dec. 2003).

To sum up, one of the core papers concludes, we “(...) do not propose standardized methodology for universal application. Rather, ... (we) consider the choice of method should be sensitive to the nature of the proposed action, the stage in the planning and project cycle at which it is being appraised and the context (institutional, regulatory, cultural) in which the appraisal will take place. (...) a ‘step-by-step’ approach to the improvement in integrated appraisal procedures and methodologies may be needed” (Bond et al. 2001 p. 1020f., cf. also George and Kirkpatrick 2003).

²³ Cf. E.g. Bond et al. 2001; George and Kirkpatrick 2003.

²⁴ 24 and 25 November 2003 in Manchester

2.5 Lessons from micro-finance?

There is a much longer tradition as well as much more experience for assessing impact in the field of micro-finance for SED. There exists a vast number of studies and publications dealing with impact assessment of micro-finance programs.²⁵ It is generally argued that financial services to SED can not be compared with the non-financial services to SED because there are only few standardised instruments for micro-finance which are contrasted with a huge variety of instruments for SED. Micro-finance is generally an umbrella term that refers to the provision of a broad range of services such as deposits, loans, payment services, money transfers and insurance to poor and low-income households and their micro-enterprises (Khawari 2004, p. 3). In a much narrower sense though, micro-finance is often referred to as micro-credit for tiny informal businesses of micro-entrepreneurs (ibid p. 4). An outstanding feature of micro-finance programs the end user of the services are by definition the poor, what is rarely the case in non-financial SED programs.

In the following three example in the field of micro-finance – one two for micro-finance particularly for small enterprises (AIMS, Imp-Act) and one for micro-finance in general (CGAP) are briefly presented.

2.5.1 *Assessing the Impact of Micro-enterprise Services (AIMS) research project*

Between 1995 to 2001 the USAID's office for Micro-enterprise Development supported a number of research and publication activities under the "Assessing the Impact of Micro-enterprise Services" (AIMS) project. This program covered primarily financial services to micro-enterprises and its impact. The studies explore salient micro-enterprise-related issues in the form of desk studies; report on mission technical assistance; discuss the development and testing of cost-effective practical tools that can be used to track and assess the impact of practitioners' micro-enterprise programs (cf. review on M+E of NGOs below); present the findings of methodologically rigorous core impact assessments (cf. below); highlight the activities of the Consultative Group to Assist the Poor (CGAP); and contribute to the World Development Report.

Within the framework of AIMS a review of *monitoring and evaluation systems of four non-government organisations* (NGOs) providing business development programs 'beyond credit alone' for micro-enterprises in development countries, was conducted by Hyman/Dearden 1998. The purpose of this study was to identify existing practices before developing new tools or indicators. Interesting findings are, all the four NGOs²⁶:

- had *hybrid systems* for monitoring output and each measured aspects of impact, while reserving more extensive impact assessment for less frequent evaluations or special studies;
- relied on both *quantitative and qualitative* studies, although the mix of method varied;
- addressed impact on client *enterprises* and at least partially also on the *households* of owners and workers and other enterprises;
- considered *broader impact* within the local economy, but this was limited to backward and forward linkages to related enterprises not directly assisted by the programs;

²⁵ For a recent review of literature on micro-finance cf. Khawari 2004.

²⁶ These were (i) Appropriate Technology International; (ii) Freedom from Hunger, (iii) Save the Children and (iv) TechnoServe.

- used a *combination* of field staff and hired external consultants to gather impact data;
- collected *longitudinal data* and used a 'standard' set of impact indicators.

With regard to indicators, each organisation had adopted or was testing its own set of indicators.²⁷ Differences in impact indicators across organisations reflected the diversity of their micro-enterprise programs and institutional mandates (cf. examples in figure 8.). Some NGOs noted the necessity of tailoring clients survey questionnaires to specific projects to gather this information, rather than applying the standard list of indicators directly (ibid p. 262). Some others had dropped some indicators that proved burdensome for field staff or did not generate useful information, such as person-days of labour time generated by projects. It was concluded that the diversity of micro-enterprise program objectives and strategies makes it difficult to establish a standard set of indicators applicable to all executing or implementing agencies (ibid. p. 274).

Figure 8: Example indicators used by NGOs for assessing impact of micro-enterprise programs

Unit of impact	Type of impact
Enterprise Level	<ul style="list-style-type: none"> - Number of participating enterprises - Number of producers - Gross income - Net income - Value Added - Change in Assets - Employment - Enterprises management practices - Financing - Sustainability of enterprises
Household Level	<ul style="list-style-type: none"> - Net income - Consumer savings - Health (nutrition, water supply, sanitation) - Educational progress - Expenditures - Housing or other material goods consumed - Empowerment (solidarity, relations in community, evidence of group self-management, etc.)
Community Level <i>Source: Hyman/Dearden 1998 Table 2, p. 265 f.</i>	<ul style="list-style-type: none"> - Environmental or natural resources impact - Policy impact (e.g. changes in regulations and policies affecting micro-enterprises) - Community infrastructure and services (e.g. access to electricity, water, sanitation etc.)

Moreover, it was found that most NGOs can only afford to do relatively small sample surveys due to the cost and time involved in data collection and analysis. As a result, evaluation surveys have mainly been confined to project participants and sample sizes commonly range from 25 to 100 respondents (ibid p. 267).

Even when stratified samples sizes were large enough to allow statistical testing, none of the organizations regularly tested the statistical significance. None of the organisations did regularly survey non-participants in program areas, but some did occasionally. Consequently, the regular impact assessment of these four NGOs did not comply with the scientific or econometric rigour, usually requested in scientific impact assessment. In spite of that all four organisations reported that impact findings often led to changes in the design of the evaluated projects. Three relied on the information for strategic planning and design of new projects, and two in resource allocation decision.

Among the new instruments developed and tested was the *core impact assessments* (CIA) as a key component of the AIMS project (cf. Dunn 2002a, 2002 b). The CIA were methodol-

²⁷ Most of these indicators are also part used in the BDS performance measurement framework cf. McVay et al. 2001, Miehlsbradt/ McMay 2003.

ogically rigorous, longitudinal impact studies of three microenterprise programs: SEWA Bank in India, Mibanco in Peru, and Zambuko Trust in Zimbabwe.²⁸

The three studies grew out of a common research design and tested the same set of core impact hypotheses using a similar research approach. In order to assess the impacts of micro-finance programs, it was found to be necessary to address at least three specific conceptual challenges: (i) fungibility, (ii) attribution and (iii) selection bias. The *household economic portfolio* (HHEP) model was considered to be the most suitable for addressing these challenges. The conceptual model widens the unit of analysis beyond the single enterprise to the entire household portfolio.²⁹ There were four major design features that contributed to the strength of the research strategy:

- “By widening the unit of analysis to include impacts on the entire household economy, the conceptual framework addressed the problem of fungibility and provided a logical framework for attributing the observed impacts to the program services received.
- The mixed-method approach combined quantitative and qualitative data to yield a much more informed view about how and why impacts occur and to strengthen the case for attribution.
- The quasi-experimental design controlled for the influence of external, non-program factors that affect the outcomes for both clients and non-clients, thus establishing the underlying trends related to the counterfactual. The fact that the non-clients in the control group were eligible for program participation helped to reduce selection bias and improve the case for attribution.
- The use of panel data, along with a statistical approach that incorporated information on the starting values of the impact variables, helped to reduce some of the influence of selection bias on the impact results” (Dunn 2002b, p. ix).

The three core impact assessments produced a relatively rigorous set of impact evaluations that point the way toward improved methodology in the field. These studies established some strong, plausible evidence for the impacts of micro-enterprise services. The studies clearly indicated that impact occurs not always where one might expect to find it.

However, it was also stated that the research approach had several important limitations, which are:

- “Selection bias may not have been entirely eliminated. The analysis of covariance (ANCOVA) approach does not control for differences on unobserved variables that may affect outcomes, possibly leading to overestimation of positive impacts and underestimation of negative impacts.
- The baseline measures might already reflect impacts because they are not true pre-treatment measures of the outcome variables, possibly leading to underestimation of both positive and negative impacts. While selection bias and the lack of a pre-treatment baseline may have opposite effects on positive impacts, they both may lead to underestimation of negative impacts.

²⁸ All three of these programs provided micro-enterprise credit, with the SEWA program also offering savings services and non-enterprise credit and the Zambuko program providing business training.

²⁹ For a more details on the HHEP and conceptual framework as such cf. AIMS-team 2001 a and 2001 b.

- The research approach relied on relatively unsophisticated measures of program participation and there were measurement weaknesses in some of the impact indicators” (ibid p. x).

These methodological limitations of the research strategy should be kept in mind when interpreting the research findings and planning future studies.

2.5.2 *The Consultative Group to Assist the Poorest -CGAP*

The Consultative Group to Assist the Poor (CGAP), is a consortium of 28 public and private development agencies working together to expand access to financial services for the poor, referred to as micro-finance. CGAP is committed to the twin objectives of increasing the financial and institutional sustainability of micro-finance institutions (MFIs) and deepening their poverty focus (i.e. increasing poverty outreach and impact on poorer people). As part of this commitment, and in its role as a service provider to the micro-finance industry, CGAP has continually endeavoured to provide tools that allow for greater transparency of MFI performance in meeting such objectives.³⁰

The focus on transparency thus far, has centred around financial performance. There has been a marked under-emphasis on the equally important issue of transparency of the poverty focus of MFIs. The CGAP Poverty Assessment Tool has therefore been developed as a much needed tool to provide transparency on the depth of poverty outreach of MFIs (cf. CGAP 2003). The tool is intended primarily for use by donors that focus on poverty outreach to form part of their appraisal of MFIs for funding.

While the tool is relatively simple it still presupposes a certain level of statistical skills and familiarity with the statistical package SPSS. It is expected therefore that the team implementing the tool have both a social science background with knowledge of poverty issues as well as basic statistical experience. Experience has shown that most countries have research institutions and private firms that are well qualified for such activity. In all the field tests concluded so far, CGAP has used local teams.³¹ Given the need for technical expertise and the resources involved in running such surveys, MFIs themselves are generally not encouraged to take on the tasks themselves.

The tool involves a survey of 200 randomly selected clients and 300 non clients, takes about four months to complete and costs around \$10,000.

The survey tries to elicit information on different dimensions of poverty. However the questionnaire has been designed to be simple and operational. Obviously, specific questions and the wording of each question has to be modified and adapted in each study to ensure that it addresses the social and cultural contexts of the region where the study takes place. The emphasis on a variety of indicators stem from a realization that poverty is multidimensional and that any one indicator would not be capable of capturing poverty levels across different countries and contexts. Additionally the use of multiple indicators provides for a better approximation of poverty levels. It is recommended to use this tool in conjunction with other

³⁰ The CGAP Appraisal Format was developed to provide practical guidelines and indicators for measuring MFI performance in governance, management and leadership, mission and plans, systems, operations, human resource management, products, portfolio quality and financial analysis.

³¹ Field tests were successfully completed in four countries.

appraisal tools (such as the CGAP Appraisal Format) to ensure a holistic understanding of MFIs.

The tool was primarily designed for those who would require a more standardized, globally applicable and rigorous set of indicators to make poverty-focused funding decisions. It claims to provide rigorous data on the levels of poverty of clients relative to people within the same community through the construction of a multidimensional poverty index that allows for comparisons between MFIs and across countries.

2.5.3 *Lessons learnt*

Despite an impressive body of studies devoted to the assessment of micro-finance programs, these are not beyond any doubts. A recent review of major studies concludes:³² “Use of diverse methodologies: Common findings but the critical issue still unsolved (...) Quasi-experimental design was adopted in each of the studies reviewed. But in most cases, no comparative analysis of the control and treatment groups was provided” (Kahlily 2004, p. 343). And the article continues with “Although the findings that emerged from the review of major studies are more or less similar, the validity of the impact remains a question since neither of the studies could tackle the issue of fungibility, and separate out the effects of informal, formal and micro-credit at the household level (ibid p. 344). “Taken as a whole, the evidence is mixed. (...) Nonetheless, the preponderance of evidence indicates the existence of a range of positive impacts on micro- enterprise development (e.g. sales, profits, and asset accumulation) and on client and household economic and social well-being. The evidence also suggests that the nature of these impact and the causal impact chain are highly variable and less straightforward than often portrayed in the promotional brochures of micro-finance institutions” (Woller and Parsons, 2002, p. 133f.) Likewise, Khawari (2004 p. 25) states: “While reviewing most of the existing literature on micro-finance, the institutions, their diverse structures, the tools and mechanisms employed, the actual running of their operations, their social and economical impacts; all seem to function somehow like advertisements. A significant part of the theoretical approaches are sustained with anecdotes from a particular set of programs rather than empirical regularities.” “While innumerable studies have been made, only a handful used appropriate control frameworks and sizeable samples that meet the agenda of answering the questions of financial sustainability, impact and outreach to the poorest and the trade off in pursuing the goals simultaneously” (ibid p. 27). This reveals the difficulty of “proving” impact even in the somewhat “easier” field of micro-finance.

The problem of comparability between the financial and non-financial services to small enterprises does exist. In micro-finance programs the end user is by definition poor whereas this is rarely the case in non-financial SED interventions. However, this difference does affect basically performance measurement but not impact assessment as such. Hence, one could definitely draw from the experiences gathered so far in the micro-finance field. With little adaptation most of these tools could be applied in the context of non-financial services to micro, small and medium enterprises.

The above examples illustrate the various stages and attempts to assess impact of micro-finance services. This supports the idea of

³² Among others the authors reviews also studies undertaken under the AIMS CIA program Kahlily 2004.

- also micro finance has not yet found the “golden pathway” to a rigorous and credible impact assessment;
- not relying on a “one-dimensional” approach to overarching development policy goals;
- using both quantitative and qualitative methods for assessing impact;
- adapting the instruments to the regional and local context;
- looking for results not only in a “narrow straightforward” causal chain of impact but also elsewhere;
- staying pragmatic and realistic with regard to what can actually be assessed
- and finally no impact assessment should be the same, unless organisations and the contexts in which they work are exactly the same.

3. Alternative ways to look at impact (assessment)

3.1 Orientation towards Results: GTZ’s New Impact Monitoring

As earlier indicated, GTZ underwent recently a radical change in its planning and monitoring philosophy. This is based on two major considerations:

1. “Typical technical cooperation projects are a small tool in relation to a task like poverty reduction. Can such projects really be expected to prove their contribution to highly aggregated development goals? (...) many technical cooperation projects make a contribution to economic and social development. But to provide accurate proof is, for the most part, impossible. Statistically the contribution cannot be isolated from a large number of other contributing factors ...” (Kuby, Jan. 2000 p. 4f).
2. Moreover, it is increasingly recognised that innovation is not a single event and not a linear process. Rather, it is considered to be a complex process over time, that depends on the interaction, influence and power relation of many different groups and actors. Innovation is the outcome of many incremental improvements that result in progress not because of a governing idea, but because of a highly recursive process of adaptation and assimilation (cf. Kuby, Jan. 2000, p. 5).

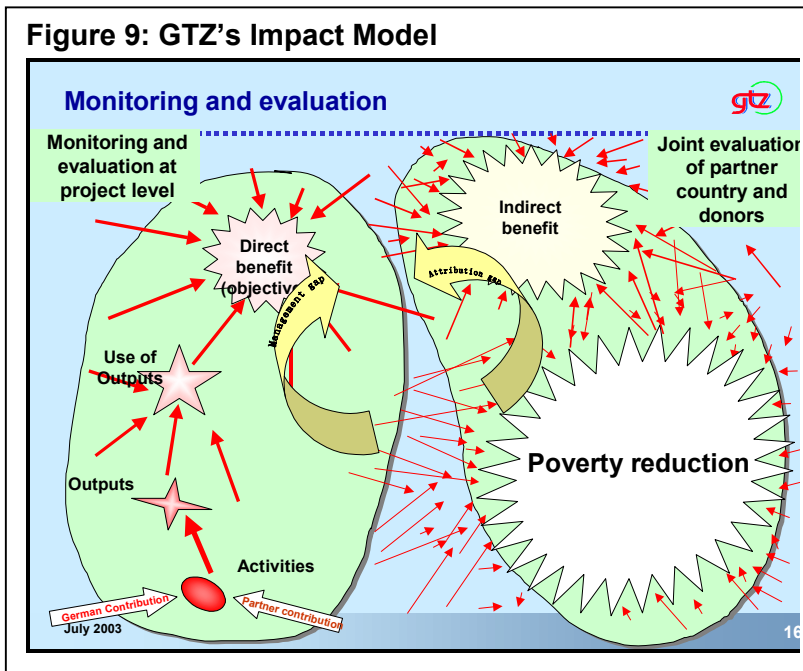
These thoughts lead away from the rigidity of the former logical framework planning (ZOPP/GOPP) and the related input-output analysis to a more open and flexible approach with an emphasis on results (cf. Kuby, Jan. 2000). Accordingly, impacts

- are changes in a status quo;
- can be intentional and unintentional, expected or unexpected, positive or negative;
- can occur from the very beginning of the intervention, during the entire project term and in different areas; and
- are the result of social interaction.

The essence of GTZ’s impact model is illustrated in figure 9 below and draws particular attention to the various benefits generated by a development intervention.³³

Having such a complex process in mind, one also has to take into account the limits of a project. GTZ distinguishes between a direct area and an indirect area of influence of a program; these are the left and right “potatoes” in the figure. The left one includes (i) activities,

³³ According to GTZ, AURA-Presentation english, 14 July 2003.



(ii) output, (iii) use of outputs and (iv) direct benefit or impact that lie logically within the influence of a project. The right one includes the indirect benefit and impact at the level of overarching development goals such as poverty reduction. These types of impact are beyond the direct and exclusive influence of a program. The influence of confounding external factors increases with the level of aggregation; and success or

failure cannot any more directly attributed to a single development intervention.

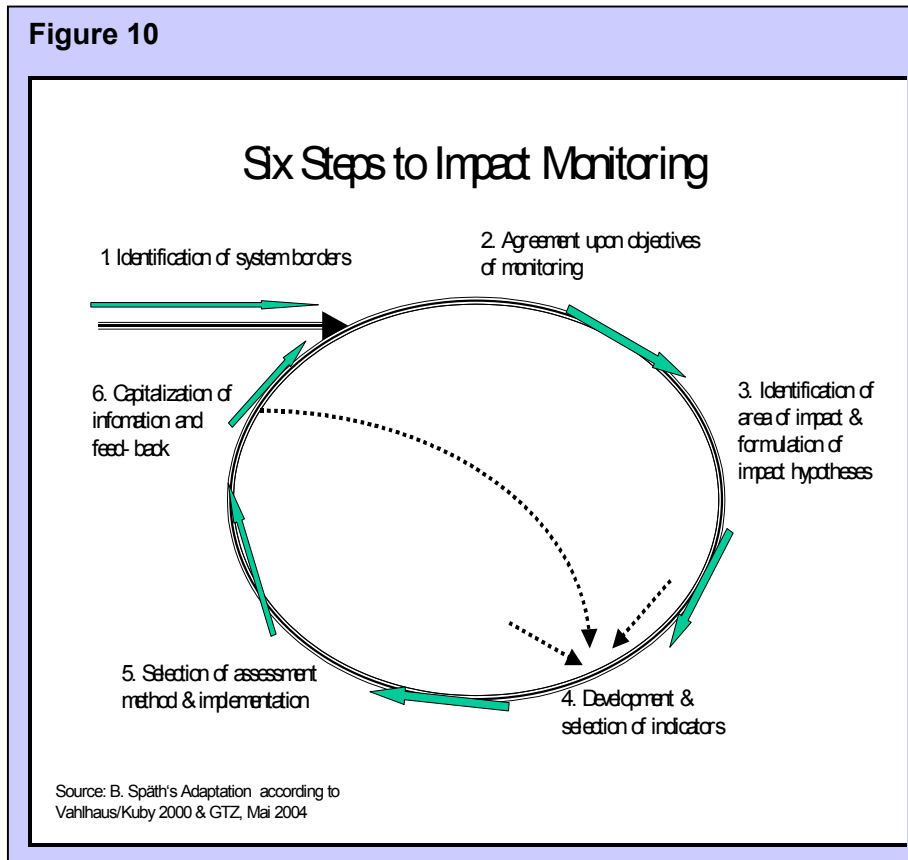
3.2 Six Steps to impact monitoring

According to the new guideline for “*impact oriented monitoring for technical cooperation projects*”, there are six steps for doing impact monitoring (GTZ, Mai 2004). The practical part of the guideline draws especially from an earlier work of Vahlhaus and Kuby (2000), who have developed “*a framework for impact monitoring of poverty alleviation in projects for economic and employment promotion*”, but also on Herweg and Steiner (2002). Accordingly, “six steps” lead the road to impact monitoring (cf. figure 10).

Quite some of the activities, which are considered to be relevant for impact monitoring, are actually part of project preparation, such as planning and design. This applies also to identifying the system borders of a project, the formulation of impact hypotheses as well as of indicators.

3.2.1. Identification of system borders

This is basically an exercise for defining and distinguishing between



the area of direct and indirect influence of a project (cf. figure 9 left and right potato). What is the project? Who – partners, target groups, and others - belongs to the stakeholders of the project? Where does the “context” start? Such questions will help to draw a border line between the areas of direct and indirect influence of a project. At the same time, it is useful to select and concentrate on “key stakeholders”.

3.2.2 *Agreement upon objectives of the monitoring*

In a second step, it is considered to be crucial to clarify the objectives of a monitoring process. This has to be done in view of different interests and expectations of those who will be involved in the monitoring. Important questions to be raised here are: Who needs which information for what purpose? At this point it should also be clear what financial, staff and time resources are available for the monitoring. Once this is sorted out the further procedure is more or less predetermined.

3.2.3 *Identification of the areas of impact & formulation of impact hypotheses*

This is basically a setting of priorities by selecting from a complex and continuous flow of information what is most relevant to the purpose of the monitoring. Out of many possibilities key areas or variables are chosen in which impact is expected. Typically, in SME-promotion program induced changes are expected at:

- the macro-level, e.g. regulatory environment, fiscal incentives;
- the meso-level, e.g. lobbying position of SME-chambers, or number and quality of BDS;
- the micro-level, e.g. inter-firm cooperation, number of SMEs (cf. Vahlhaus/Kuby 2000, Part II, p. 16).

This may be the place to decide what units of impact should be looked at. Is it going to be at the level of institution /organisation, enterprise, household or individuals. Additionally one can consider to look at different social, geographic, political or administrative units such as ethnic community, local or regional level, and municipal, provincial or state level.

Depending on the type of program one can select areas of change at the level of “soft skills” or behaviour of beneficiaries such as knowledge, ability, perception and interaction.

At this point it is suggested to verify and adapt existing impact hypotheses, and if necessary to formulate new ones. An impact hypothesis is considered to be an assumption about cause and effect relations, that is either based on experience or even proved by scientific research.³⁴

3.2.4 *Development and selection of indicators*

Indicators are considered to be part of the design and planning of a project. They are tools to further clarify and more precisely define the objectives of a project and its desired impact upon the intended beneficiaries. At the same time indicators are milestones or landmarks which help to observe the progress and changes of a project. The indicators have to be specified in terms of quantity, quality, time, geographic location/outreach, target groups or beneficiaries: How much? In what way? or How well? When? Where? With or for whom? In

³⁴ For more information cf. Vahlhaus/Kuby 2000 Part II, p. 15 ff. the selection of fields of impact as well as the formulation of impact hypothesis for different project types (financial services, SME-promotion, vocational training, policy advice) is treated in more detail.

addition, one would have to specify reference points, that clearly indicate the observed changes. One could use data from a baseline study, data of comparable situations or a certain quality standard.³⁵

3.2.5 Selection of assessment method & implementation

The selection of an assessment method depends greatly on the type and quality of data / information needed, as well as on the resources (financial, personal, time) available. Moreover, one also has to decide on who is going to make the assessment (internal, external) as well as on how often. In any case, one will have to consider both costs and benefits when opting for a specific monitoring method.³⁶

3.2.6 Capitalization of information and feed-back

Finally, the valuable has to be extracted out of the gathered data and information. In a first step there would be a systematic evaluation of the information available and its interpretation by the parties concerned. In a second step, the results will have to be fed-back into the "system" in order to allow adaptation in the sense of "lessons learnt".

3.2.7 Summing up

The GTZ guideline does not prefer a specific assessment method. For example group discussions, participatory observation, action research, participatory rural appraisal are considered as valuable assessment tools. Though, these instruments render (subjective) information based on the perception of people, in most cases they are considered to be more viable as merely statistical data. The major criteria for selecting an approach are on the one hand the purpose of assessment, and on the other the expected costs and benefits.

As an executing agency GTZ takes only responsibility jointly with the project partner for the direct impact of a development intervention (left "potato" in figure 9). Nevertheless it is expected that the monitoring gives an idea on whether the project is on

Figure 11: Do's and Don'ts" for Impact Monitoring with a Poverty Focus in Policy Advisory Projects

Do orient the project work towards impact at all stages of the project cycle.

Do include the view on poverty.

Don't have unrealistic expectations of the impact monitoring you will perform in your project. In most cases, you will not be able to *prove* your project had significant impact. Don't expect social science research or scientifically rigorous, "objective" results. Don't even expect the "Truth" (although you should always aim for honesty).

Do abandon the logframe or PPM "mindset". Impact monitoring provides the opportunity to perform a reality check of the project concept and plan. This cannot be done if you refer to the same picture of the project as that provided by the logframe.

Don't automatically generate impact indicators for all "Results" on the PPM. This would lead to a rigid, mechanistic view of impact and would undermine the main advantage of the orientation towards impact – that it provides a "living" picture of the effects of the project and grounds the project work in reality.

Do develop instruments for observing impact and ways to institutionalize impact monitoring *together* with the stakeholders.

Don't base the impact monitoring system or instruments on assumptions that you will produce cost/benefit analysis or some sort of rating on an objective "impact scale" – only very few projects are of a type that the impact assessment will lead automatically to these kinds of comparison.

Do embrace an attitude of learning.

Don't forget: in monitoring the impact of policy advisory projects, the process is at least as important as the findings.

Do rest assured that, as long as you invest honest effort into it, the impact monitoring your project performs will add to the body of knowledge on development projects and on impact assessment.

Do be bold, be creative. Just do it!

Source: Lobb-Rabe, 2000, p. 69

³⁵ Cf. again for more information especially for economic and employment promotion projects cf. Vahlhaus/Kuby 2000 Part II, p. 27ff.

³⁶ For methods and instruments with a particular view on economic and employment promotion projects cf. again Vahlhaus/Kuby 2000 Part II, p. 64 ff.; for handy-craft projects cf. Gomm 2001, for policy advisory programs with a poverty focus cf. Lobb-Rabe 2001

track (target corridor) to contribute to higher development goals (right potato). Impact assessment of higher development policy goals is, however, considered to be a joint task of the donor(s) and the partner country. This is of particularly relevant for monitoring the progress in the frame of the PRSP and MDGs (cf. annex 2).

According to these documents the limits of impact assessment and monitoring seem to be quite clear, which is also emphasized by the recommendation on ‘do’s and don’ts for impact’ in figure 11. It is important to note, that for the year 2004, GTZ has set an “orientation towards development results” as a strategic goal. So we can expect to get more practical information in due time.

3.3 A participatory learning impact assessment method - Method for Impact Assessment of Poverty Alleviation Projects (MAPP)

The ‘Method for Impact Assessment of Projects and Programs’ (MAPP)³⁷ is a methodological framework combining a qualitative approach with participatory assessment instruments, but, it includes also a quantification step. It orients itself towards principles and procedures of Participatory Rural Appraisal (PRA) methodology, such as:

- (i) *triangulation*: the collection of distinct data with different tools in order to proof or raise the validity of the data;
- (ii) *optimal ignorance*: the capability to select relevant data and to avoid an information overkill;
- (iii) *communal learning*: the findings of an assessment are the result of a communication process among relevant groups.

A major element of this method is conducting of workshops with representatives of relevant stakeholders. Perceived key processes are jointly reflected in structured group discussions, in which at least six interlinked and logically connected steps are accomplished. These are (i) lifeline, (ii) trend analysis, (iii) activity list, (iv) influence matrix, (v) transect – or data cross checking and (vi) development and impact profile.

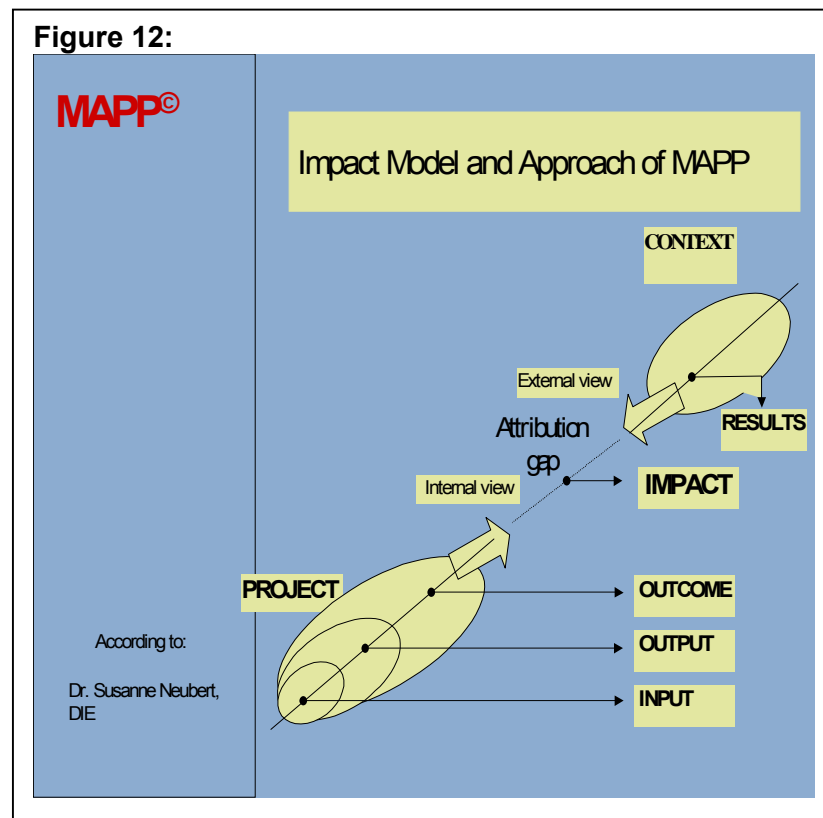
Though, originally developed for social impact analysis of resource management projects in Sub-Saharan Africa, MAPP can also be applied to other types of programs and for other regions. The method is particularly suited for analysing more complex development goals that can usually not be evaluated with the help of one or more quantitative indicators. Such multi-dimensional goals are for example poverty reduction, democratization, good governance, economic and sustainable development. They can only be described in a more comprehensive way by using three or four key criteria or processes. The definition of the respective key criteria is the outcome of a joint process of analysis. At the same time this assures the adaptation of the analysis to each specific local and thematic characteristics.

For example, in order to appraise poverty impact, MAPP draws from the multidimensional definition of poverty developed by Amartya Sen (1999). His ideas are operationalized in MAPP by the concept of the four key social processes, which are the evolution of (i) livelihood, (ii) access to resources, (iii) knowledge and (iv) rights. This understanding of poverty goes clearly beyond a narrow definition of poverty that is based on a change of income of the individual or household as suggested earlier by Oldsman or the indicators of the first target of

³⁷ The method was developed by Susanne Neubert, staff member of the German Development Institute (DIE/GDI), Bonn.

the Millennium Development Goals (cf. figure 4). In fact, the method claims to be able to assess whether poverty-reducing processes took actually place at the level of the target groups. The program benefits can even be differentiated according to different social (target) groups, such as gender specific, socio-professional etc.

In the conventional approach the project and usually some sort of input-output analysis is in the centre of the assessment. In contrast to that, MAPP considers the context of a program as a point of departure (cf. figure 12). With the help of a sequence of specific tools, the perception of the development process in the daily life of the population is evaluated. Only in a second step the observed changes are attributed to distinct project measures, initiatives of the population itself, intervention of other donors or external factors.



It is not necessary to define detailed chains of impact or causalities beforehand. Although, impact hypotheses are considered as relevant for the design and planning of projects. By going away from a narrow project-oriented approach – including logical frameworks, planning matrices, pre-formulated chains of impact and causalities – the view widens for the changes that have actually taken place within a given period of observation. This allows also for assessing unintended positive and negative effects. Thus, this method goes clearly beyond a narrow context of result monitoring and /or performance measurement.

A representative coverage of a program is obtained through the representation of (i) different social groups (stakeholders) and (ii) different branches, regions or administrative units at the workshop. In most cases several workshops will be conducted e.g. in project regions where the performance is particular good, bad or mediocre, or where activities have just started, are in the middle of the road or are being consolidated. Presently, a coverage of 10% of those people affected by the program is considered as ideal. Since representatives of the different groups take part in the workshop mutual control or validation of individual statements and assessments is an inbuilt part of the evaluation process.

The method can only be applied where implementation has taken place, that means something has actually happened and has caused changes. Without problems the method can easily be applied to projects that are in their tenth year of implementation.

The existence of baseline surveys and studies is not necessary, however, it would be useful for final cross checking of data and information. In a final comparison and interpretation of results with other data and information (such as existing census data, program monitoring, expert opinions) a comprehensive understanding can be obtained of (i) what developmental processes took actually place; (ii) what did the project contribute to it; (iii) what was the role of other donors or local initiatives; and (iv) how did changes in the overall framework influence this development.

Accordingly, one could judge on whether the project's activities are in the "target corridor" for achieving the overarching development goals. In addition, by combining and interpreting the different information from various sources, it would be possible to reduce the "attribution" gap between project input and aggregated results.

To sum up, MAPP maintains:

- to be a workable assessment method, with a modest requirement of time, as well as financial and human resources;
- to provide an optimum of information value, transparency and validity;
- to allow a comparing of results.

4. Concluding remarks

For the concluding remarks we should recall the opening citation, of impact assessment being as much an art as a science. It continues:

"Enhancing the contribution that impact assessment can make to development goals requires both better science and better art. The scientific improvements relate to improving standards of measurement, sampling and analytical technique. Econometricians and statisticians are particularly concerned with this field. Improving the "art" of impact assessment has at least three strands. One concerns making more systematic and informed judgements about the overall design of IAs in relation to their costs, specific objectives and contexts. The second is about what mixes of impact assessment methods are most appropriate for any given study. The third relates to increasing our understanding of the ways in which the results of IA studies influence policy-makers and (...) managers" (Hulme 2000, p. 79f.).

There exists a vast and diverse body of theoretical literature as well as numerous studies and research results on impact assessment. It is beyond the scope of this paper to cover the bulk of existing literature, but hopefully the most relevant and productive ones were included. Only recently some significant publications were made available to me that could not yet fully be covered within the frame of this paper, these are:

- Nadvi's and Barrientos' development of *a methodology for poverty and social impact assessment of cluster development initiatives* (2004)
- A special issue of the IDS Bulletin on *Microfinance, Poverty and Social Performance*, October 2003
- A special issue of the Journal of International Development on *The Social Impact of Microfinance*, April 2004.

In fact, it is an ongoing process and like all innovation, development of methodologies, tools, instruments and protocols cannot be viewed as an entirely rational process: Semantics, rela-

tionships, resources and power are important (cf. Wright and Copestake 2004, p. 357). There is a wealth of experience with impact assessment methods: good, bad and ugly. Therefore it is important to locate any new proposal in the context of already crowded and confused quasi-market place (cf. *ibid*, p. 358). For any new proposal it is not as much a question of whether it constitutes truly a new one, but rather, whether it helps to cope in an appropriate way with the ever increasing demand for impact assessment. In the following, six salient points of impact assessment are once pointed out: (i) assessment of higher development policy goals, (ii) causality and attribution, (iii) plausibility, (iv) experience and scientific research, (v) concentration on the feasible and (iv) a common understanding.

(i) assessment of higher development policy goals

In the previous chapters it was underlined, that most overarching development goals – such as poverty reduction, economic development, democracy, sustainability, - are multi-dimensional.³⁸ They can only be described or defined by a number of key criteria. Hence, such a goal cannot adequately be captured by looking only at one dimension or perhaps one or two quantitative indicators. A methodological approach based on a single quantitative aspect e.g. a change in income, proposed by Oldsman, can never grasp the complexity of the issue. It is therefore suggested to turn to approaches and methods that allow to draw a more comprehensive and social complex picture of impact and its related processes, such as for example the methodological framework of “MAPP”.

(ii) problems of causality and attribution³⁹

Although, most would appreciate greatly getting exact data on impact, one has to be realistic about the limits of econometric and statistical methods. “... (T)he ability of most programs to reduce poverty in a particular area is likely to be limited given the scale of the problem and the relative magnitude of the intervention (i.e., the effect size of programs is likely to be small). This means that it will be difficult to identify an impact even if it exists. Large datasets – which are likely to be costly to compile – will be needed to estimate poverty impacts” (Oldsman 2003 p. 10). Moreover, to get valid data would imply that the net effect of an intervention is isolated from all other extraneous or confounding factors that influence defined impact. However, we have learned that many (confounding) factors are at play in effecting the changes observed at a highly-aggregated level. Consequently, these changes cannot be reliably and proportionately traced back to any one intervention or program. Instead of trying to circumvent the methodological problem of attribution at a higher level of impact assessment with onerous statistical methods with doubtful outcome, one should rather turn to identify plausible explanation for what one can observe at field level.

(iii) reaching plausibility⁴⁰

Impacts of a project are often unforeseen. Therefore the crucial question is not only whether planned results have been achieved, but **what** results have been achieved, and whether these results **effect** the desired changes in the live of people, a region or a country. A GTZ working group concluded: “...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. (...) the reasonable compromise is to

³⁸ Cf. particularly chapter 2.3.7, 2.5.2 and 3.4.

³⁹ Cf. particularly chapter 1.4.1, 2.3.7, and 3.1

⁴⁰ Cf. this is particularly underlined by the GTZ-approach under 3.1 – 3.3 and MAPP 3.4

establish plausible links between the development efforts and the observed changes” (2001 p. 1). And it continues to stress “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” (ibid p. 3). This is considered to be a reasonable compromise because a) it only requires an adequate amount of resources and b) it delivers appropriate information for the purpose of learning, managing and accountability.

(iv) relying on scientific research and experience⁴¹

It is not necessary that each donor and / or program will have to conduct in depth research on small enterprise development. First of all, there exists a lot of scholarly literature on micro, small and medium enterprises dealing with various aspects of their role in economic and social development. After all, with the promulgated the informal sector concept by the ILO in 1972 many studies and research programmes were conducted in the wider field of small enterprise development (cf. Späth 1997). E.g. a more recent study made by a team from the World Bank analysed the relationship between SMEs, growth, development and poverty (cf. Beck, Thorsten et al. 2003). Secondly, a good documentation of program implementation as well as case studies of the success and failure of SED interventions would add to the body of knowledge and experience, from which one could draw for further learning. Capitalization of existing knowledge and experience could lay the foundation for developing plausible impact hypotheses.

(v) concentration on the feasible

In the “usual business” of development cooperation quite some progress could be achieved, if:

- feasibility studies and baseline surveys would be made properly before or the latest at the beginning of a program or at a new conceptual phase,
- monitoring would go beyond the mere input – output analysis, but would also cover aspects of performance measurement;
- a focus on assessing direct program impact would be part of regular program reviews and evaluations.

A good documentation based on a continuous and/or regular collection of information would already satisfy most requirements for program management and steering as well as for learning and improving. Although, one may have diverging ideas about the type of data to be collected and the methodology to be used, in principle this idea of a more continuous collection of information at lower levels of impact is supported by Oldsman: “Greater attention should be placed on conducting rigorous evaluations of programs in terms of their impact on firm-level performance using experimental or quasi-experimental designs. As noted above, these impacts are necessary in order to achieve the ultimate goal of poverty reduction. Evaluations of intermediate objectives would require the development of good firm-level data based on longitudinal surveys” (ibid. 2003, p. 20). This would additionally provide a solid base for underpinning plausibly the relationship between program interventions and impact at higher levels.

⁴¹ Cf. this is particularly underlined by the GTZ-approach under 3.1 – 3.3.

(iv) a common understanding

What should have become clear in discussing all these different studies and documents, there exists not a single and only way for doing impact assessment.⁴² Every impact assessment has a conceptual framework at its centre. At least would have to do a contextual adaptation of assessments frameworks. It is therefore recommended, that those who are involved in an impact assessment, should check first whether there is a common understanding of what is going to be assessed in terms of terminology, level, unit and type of impact. The check list of questions in Annex 3 could be a useful instrument for such a clarification process. This is not only relevant for avoiding misunderstanding and frustration it is especially crucial for selecting an assessment approach and/or methodology. Such a common understanding is also necessary if results are to be comparable between different phases of a program, between the same type of programs but in different countries, between different types of programs and between programs of different donors.

*** BS 02.08.04 ***

⁴² This is underlined by the outcome of a virtual discussion of Imp-Act, a network for micro-finance impact assessment, which identified the greatest danger to impact assessment in conformity. The virtual meeting took place in April 2001 and was documented by Simanowitz (2001) cf. also www.imp-act.org.

Annex 1: Websites of Interest on Impact Assessment

Réseau IMPACT – Inégalité / Micro-macro / Pauvreté / Acteurs Network IMPACT – Inequalities / Micro Macro / Poverty / Actors

This network does not explicitly address the issue of micro, small and medium enterprise development, rather it deals with impact assessment and poverty as such. **Objectives and Procedures:** The Network acts as a link between policies and strategies, on the one hand, and practices and results on the other. More specifically, the Network works on improving methods designed to facilitate the implementation of strategies for poverty alleviation and the reduction of inequalities.

It is the Network's ambition to combine experience and knowledge from different sources and create synergies among the various disciplines and professions. This will enable it to guide political choices, justify research orientations, influence the content of development-related professions, and contribute to the renewal of international cooperation.

The Network publishes notes, articles, books and reports; it organises meetings, workshops and training sessions; it also coordinates investigation missions and initiates research projects. Its work reflects a dual approach, aiding partner countries and supporting French public policy on cooperation.

Network activities are now focused on **four key theme**:

1. The links between growth, poverty and inequalities
2. The contribution of an approach based on "intermediate public policies"
3. The importance of the local dimension and decentralised authorities in poverty alleviation and the reduction of inequalities
4. Indicators, systems and impact measurement related to policies for poverty alleviation and the reduction of inequalities

www.gret.org/anim_reseau/impact.htm or www.reseau-impact.org

Germany WiRAM - “Economic Reform and Market Development“

The website of the joint BMZ and GTZ thematic programme “Economic Reform and Market Development” Besides a lot of information on the thematic programme, the website includes also information on monitoring and evaluation, and impact assessment. The content of the site is, however, mostly in German.

- Concepts and manuals
- Instruments and Methods
- Case Studies
- M&E and Poverty
- Resources

<http://www.wiram.de/wirkung/wirkung.htm>

ILO Interagency Development Knowledge Website

This interagency development knowledge website is administered by the ILO on behalf of the donor committee on small enterprises development. The mission /vision of the website is the inter-agency information exchange in the field of developing markets for business development services (BDS). It contains a variety of documents under various headings:

- Market assessment
- Project design
- Implementation
- Impact assessment
- Final documentation

<http://www.ilo.org/dyn/bds/bdssearch.home>

EDIAIS Enterprise Development Impact Assessment

The Enterprise Development Impact Assessment Information Service (EDIAIS), which is jointly managed on behalf of DFID/UK by the Institute for Development Policy Management at the University of

Manchester (IDPM), and by Women in Sustainable Enterprise (WISE) Development Ltd., website includes:

- application guidance on undertaking impact assessments in specific types of enterprise,
- toolbox items - impact assessment methods and tools
- case studies - demonstrating impact assessments carried out across a range of projects,
- EDIAIS bibliographic database

<http://www.enterprise-impact.org.uk/>

Imp-Act: Improving the Impact of Microfinance on Poverty: Action Research Programme

What is *Imp-Act*? *Imp-Act* is a global action-research programme designed to improve the quality of microfinance services and their impact on poverty by developing impact assessment systems.

Imp-Act promotes credible and useful impact assessment, building on the priorities and agendas of Microfinance Institutions (MFIs) and their clients.

Imp-Act helps institutions to develop their own learning systems, both to inform internal decision-making and to satisfy the requirements of external stakeholders.

Imp-Act works with 30 MFIs in 20 countries world-wide. It brings together academics from three UK universities, policy-makers, international NGOs and donors.

<http://www.ids.ac.uk/impact/indexguide.html>

Impact Assessment Centre: A joint *Imp-Act* and CGAP initiative

This Centre provides resources and guidance useful to practitioners and others in:

- identifying how they might usefully include impact assessment in their work,
- understanding the choices they need to make in order to get the most out of impact assessment,

engaging effectively with donors, researchers and consultants in determining the most appropriate form of IA for their organisation and stakeholder needs. The Centre is founded on three premises.

- That impact assessment is of fundamental importance for MFIs that seek to bring about positive changes in the lives of the poor, and to minimise the negative impacts that they might have.
- That impact assessment is a useful and cost-effective activity that can have a central role in the management of MFI effectiveness, in product design and evaluation, and ultimately to the overall effectiveness and efficiency of the organisation.
- That impact assessment can be done in different ways depending on the type of information that is required and for what purpose it is needed. There is no blue-print approach to impact assessment, and there is a place both for simple client monitoring systems and for "high-end" studies that seek to prove causality.

www.microfinancegateway.org/impact/index.htm

USAID – Micro Enterprise Development Programme

Besides information on the microenterprise development activities it general it contains also information of the previous AIMS-Programme, and quite a body of documents on impact assessment particularly of micro-finance projects, cf. 2. website.

<http://www.usaidmicro.org/> or <http://www.usaidmicro.org/pubs/aims/>

SEEP – The Small Enterprise Education and Promotion Network

Mission: The Small Enterprise Education and Promotion (SEEP) Network is an association of more than 50 North American private and voluntary organizations which support micro and small enterprise programs in the developing world. The Network's mission is to advance the practice of small and micro enterprise development among these organizations, their international partners, and other practitioners. This site contains also some the documentation on the work of the Performance Measurement Framework (PMF) for BDS programs: www.seepnetwork.org/bdsguide.html

Annex 2: Overarching development goals, MDGs, CDF, PRSP, PSIA etc.

Over the last years there has been an increasing list of international conventions, agreements and commitments, whereas at the same time at the national level development cooperation is increasingly questioned by the public. Donors are usually committed to a number of overarching development goals, which ultimately justify development cooperation in general as well as more specifically in a distinct sector or country. Programmes and projects will have to comply in one way or the other with these higher development goals.

Millennium Development Goals - MDGs

Box 1: The eight MDGs are:

1. Eradicate extreme poverty and hunger
2. Achieve universal primary education
3. Promote gender equality and empower women
4. Reduce child mortality
5. Improve maternal health
6. Combat HIV/AIDS, malaria, and other diseases
7. Ensure environmental sustainability
8. Develop a global partnership for development.

Source: www.undp.org/mdg

All donors have committed themselves to the UN Millennium Development Goals (MDGs). Based on the declaration of the September 2000 Millennium Summit, the International Monetary Fund (IMF), the Organization for Economic Co-operation and Development (OECD), the United Nations, and the World Bank have devised a comprehensive set of eight goals; 18 numerical targets; and over 40 quantifiable indicators to assess progress were formulated. While each goal is important in its own right, they should be viewed together as they are mutually reinforcing. Achieving them will require building capacity for effective, democratic, and accountable governance, protection of human rights, and respect for the rule of law. A good website for observing what is happening in the field of MDGs is the one on **Capacity Development for MDGs**, <http://topics.developmentgateway.org/mdg>.

CDF, PRSP and PSIA

The MDGs, particularly goal number one on poverty eradication, have given impetus to a whole set of development initiatives. The Comprehensive Development Framework (CDF) was introduced in 1999 by the World Bank. It brings together, in a unified framework, four principles to improve the effectiveness of development assistance in reducing poverty: (i) a long-term, holistic framework, (ii) results orientation; (iii) country ownership; and (iv) country-led partnership. The experience of the CDF, served as a base for the Poverty Reduction Strategy Paper (PRSP) development (cf. World Bank 2003). According to Wolfensohn, President of the World Bank Group:⁴³ "The CDF and the PRSP are the way forward to enhance country ownership and the achievement of the Millennium Development Goals."

In other words it was stated in a joint UNDP/World Bank paper: "In essence, the PRSP is the 'national roadmap' for reaching long-term MDG targets through short/medium-term policy reforms and budget restructuring."⁴⁴ PRSP describe a country's macroeconomic, structural and social policies and programs to promote growth and reduce poverty, as well as associated external financing needs.⁴⁵ PRSP are prepared by governments through a participatory process involving civil society and development partners, including the World Bank and the International Monetary Fund (IMF).

On its turn PRSP, with the strong focus on monitoring and evaluating results gave rise to the poverty and social impact analysis (PSIA). PSIA implies an analysis of the distributional impact of policy reforms on the well-being or welfare of different stakeholder groups, with particular focus on the poor

⁴³ For more information cf. web.worldbank.org/WBSITE/EXTERNAL/PROJECTS/STRATEGIES/CDF/

⁴⁴ UNDP/World Bank, How do the Millennium Development Goals Relate to Poverty Reduction Strategy Papers? Nov. 2002, <http://www.undg.org/mdgnet>

⁴⁵ cf. www.worldbank.org/poverty/strategies/index.htm

and vulnerable.⁴⁶ It is important to note, that PSIA methodology includes economic tools as well as participatory poverty and stakeholders analysis. PSIA has an important role to play in the elaboration and implementation of poverty reduction strategies in developing countries.

SDC's overarching goals

SDC's guiding principles are listed in the box to the right. They are relevant to all fields of development assistance, although the relevance will vary according to the task. These principles are also reflected in the medium-term orientation of SDC's employment and income section (E+I). There, specific references are made to poverty reduction, employment and income generation, market orientation, empowerment, gender, ecology, rural-urban development as well as capacity building.

Box 3: SDC's Mandate

Poverty alleviation
 Ensuring sustainability
 Aiding victims, resolving conflict, and ensuring peace
 Empowering the disadvantaged
 Providing access to information
 Ensuring equitable development for women and men
 Fostering understanding between cultures
 Supporting human rights, the rule of law, and democracy
 Source: *Guiding principles SDC*

⁴⁶ cf. www.worldbank.org/poverty/psia/index.htm

Annex 3: Checklist for preparing an impact assessment

What is the purpose of the assessment? Why is it done?	<ul style="list-style-type: none"> ▪ external or advocacy ▪ management / steering ▪ improving methodologies ▪ internal learning ▪ coherence between stakeholders ▪ empowerment of target groups, etc.
Who are the users? For whom is it done?	<ul style="list-style-type: none"> ▪ donor ▪ research scholars ▪ executing agency ▪ project management in the partner country ▪ project partners / stakeholders ▪ beneficiaries / target groups, etc
What impact?	<ul style="list-style-type: none"> ▪ immediate / direct impact ▪ indirect impact ▪ higher development policy impact
What type of impact?	<ul style="list-style-type: none"> ▪ economic ▪ social ▪ behavioural ▪ socio-political ▪ environmental, etc.
What unit of assessment?	<ul style="list-style-type: none"> ▪ enterprise / individual / household ▪ micro / meso / macro ▪ community / provincial / state ▪ local / regional / national level ▪ at the level of institutions, ▪ socio-economic or ethnical groups, etc.
In what context?	<ul style="list-style-type: none"> ▪ socio-economic situation ▪ political situation ▪ legal and regulatory situation, etc.
Who does it?	<ul style="list-style-type: none"> ▪ external / internal / blend
When and how often?	<ul style="list-style-type: none"> ▪ ex ante / ex post / during ▪ one time exercise <i>versus</i> regularly ▪ continuous monitoring
What resources are available	<ul style="list-style-type: none"> ▪ finance ▪ staff (quality) ▪ time
What methodology used?	quantitative / qualitative / participatory / a blend of it
What information / data is available?	<ul style="list-style-type: none"> ▪ program monitoring / performance measurement ▪ data collection by national bodies / banks ▪ other donors
Which instruments and tools to use?	<ul style="list-style-type: none"> ▪ Out of a multitude of tools and instruments for each methodology one has to select the most suitable one
When, how and to whom will there be a feed back loop?	

RELEVANT PAPERS AND ARTICLES

AIMS-Team: Conceptual Framework for Assessing the Impacts of Microenterprise Services, USAID, AIMS-Brief no. 31, December 2001 a.

AIMS-Team: Conceptual Framework for Assessing the Impacts of Microenterprise Services, Office of Microenterprise Development Global Bureau, USAID, December 2001b.

Ayagari, Meghana, Thorsten Beck, and Asli Demirgüç-Kunt, Small and Medium Size Enterprises across the Globe: A New Database, World Bank, August 26, 2003
<http://econ.worldbank.org/programs/finance/topic/sme/>

Baker, Judy L. :Evaluating the impact of development projects on poverty : a handbook for practitioners. Washington D.C.: World Bank, 2000.

Barrientos, Stephanie and Khalid Nadvi: Industrial Clusters and Poverty Reduction. Towards a methodology for poverty and social impact assessment of cluster development initiatives. A report prepared by the Institute of Development Studies University of Sussex, in cooperation with the staff of UNIDO's Small and Medium Enterprise Branch, Vienna: UNIDO, 2004.

Barry, Saliou: Erfahrungen mit dem 'System Wirkungsanalyse armutsorientierter Projekte' SWAP in der Selbsthilfeförderung des guineischen Handwerks, in: Gomm (Hrsg.) 2001, pp. 75-85.

BDS for small enterprises. Guidelines for Donor Intervention, Donor Committee for SED, 2000, often referred to as the 'blue book'. (www.sedonors.org/html/bds_guidelines.html)

Beck, Thorsten; Asli Demirguc-Kunt and Ross Levine: The Impact of SMEs on Growth, Development, and Poverty: Cross-Country Evidence, World Bank, Second Draft, September 25, 2003.
<http://econ.worldbank.org/programs/finance/topic/sme/>

BMZ: Monitoring Instruments in German Development Co-operation, Bonn 2003, summary

Bond, Richard; Johanna Curran, Colin Kirkpatrick, Norman Lee and Paul Francis: Integrated Impact Assessment for Sustainable Development: A case study approach, in: World Development, Vol. 29, 2001, No. 6, pp. 1011-1024.

CGAP: Assessing the relative poverty of microfinance clients: A CGAP operational tool, CGAP Technical Tool No. 5, Sept. 2003, downloaded from www.cgap.org

Deaton, Angus: Measuring poverty, Princeton University, Research Program in Development Studies, Mimeo, January 2003, revised July 2004.

Dollar, David and Aart Kraay: Growth *is* Good for the Poor, in Journal of Economic Growth Vol. 7, No. 3, 2002 pp. 195-225

Dunn, Elisabeth: Research strategy for the AIMS core impact assessments, USAID, AIMS Brief no. 32, March 2002a .

Dunn, Elisabeth: Research strategy for the AIMS core impact assessments, USAID, AIMS, research paper, March 2002b, 45p.

EDIAIS: Basic impact assessment at project level, often referred to as 'Core Text', mimeo, 36p. prepared Colin Kirkpatrick and Hulme, David, with contributions from Linda Mayoux, Caroline Pinder, Tertia Gavin and Clive George, by download from (www.enterprise-impact.org.uk/)

Enterprise Impact News, New Direction Conference, Issue 26, December 2003, (EDIAIS newsletter download under www.enterprise-impact.org.uk/).

Gagl, Dieter: Wirkungsbeobachtung am Beispiel der SWAP Analyse, in: Gomm (Hrsg.) 2001, pp. 61-74.

George, Clive and Colin Kirkpatrick: A practical guide to strategic impact assessment for enterprise development, Impact Assessment Research Centre, Institute for Development Policy and Management, University of Manchester, UK, October 2003.

Gibsons, Allan: Developing indicators in small enterprise development projects. A tool for people involved in designing, implementing and evaluating SED projects. SED working paper No. 1, April 2001

GTZ: Wirkungsorientiertes Monitoring. Leitfaden für Vorhaben der Technischen Zusammenarbeit, GTZ Stabstelle 04: Unternehmensentwicklung, OE 042 Interne Evaluierung, Eschborn, May 2004.

- GTZ working group 2001 (Assessing the Impact of Agricultural Research in Development): Establishing Plausibility in Impact Assessment, GTZ, Eschborn, December 2001, mimeo
- Gomm, Anja (Hrsg.): Projekterfahrungen zum Wirkungsmonitoring in der Selbsthilfe- und Handwerksförderung, GTZ Eschborn, 2001.
- Hallberg, Kris and Eric Oldsman: Framework for Evaluating the Impact of Enterprise Initiatives, SED working paper No. 3, July 2002 (www.intercooperation.ch/sed/products).
- Herweg, Karl and Kurt. Steiner: Impact Monitoring & Assessment. Bern 2002, Centre for Development and Environment (CDE, Switzerland), Vol. 1 and 2.
- Hulme, David.: Impact Assessment Methodologies for Microfinance: Theory, Experience and Better Practice, in: World Development Vol. 28, 2000, No. 1, pp. 79-98.
- Hymann, Eric L. and Kirk Dearden: Comprehensive Impact Assessment System for NGO Microenterprise Development Programs, in: World Development, Vol. 26, 1998, No. 2, pp. 261-276.
- Jarvis, Robin; James Curran, John Kitching and Geoffrey Lightfoot: The use of quantitative and qualitative criteria in the measurement of performance in small firms, in: Journal of Small Business and Enterprise Development, Vol. 7 – 2000 – No. 2, pp. 123-134.
- Khalily, Baqui: Quantitative Approach to impact analysis of microfinance programmes in Bangladesh – what have we learned? In: Journal of International Development, Vol. 16 2004, Issue 3, April, pp. 331-353.
- Khawari, Aliya: Microfinance: Does it holds its promises? A survey of literature. Hamburg: Hamburges Welt-Wirtschafts-Archiv (HWWA), HWWA Discussion Paper, Nr. 276, 2004.
- Kuby, Thomas: Turning attention towards results. How GTZ is building up its impact evaluation capacity, GTZ, mimeo, Jan. 2000.
- Lobb-Rabe, Alison: Impact Assessment with a Poverty Focus in Policy Advisory Projects: Concepts, Questions and Cases, GTZ, OE 42, Eschborn 2000
- Lübker, Malte; Graham Smith and John Weeks: **Growth and the Poor**: A comment on Dollar and Kraay, in: Journal of International Development, 14, 2002; pp. 555-571.
- Mayoux, Linda; Jahnvi Andharia, Neeta Hardikar, Sumitra Thacker and Sejal Dand: Participatory action learning in practice. Experience of a rapid participatory review of anandi, India: Enterprise Impact News, EDIAIS, Issue 31 June 2004
- Mayoux, Linda: Empowering enquirey: A new approach to investigation, EDIAIS
- McVay, Mary; Alexandra Overy Miehlabradt, and Tim Canedo: Performance Measurement Framework for Business Development Services: Technical Note on the Research Findings of the Performance Measurement Framework Field Research supported by U.S. Agency for International Development Office of Microenterprise Development, September 2001
- Microfinance, Poverty and Social Performance*. Special Issues. IDS Bulletin, Vol. 34, Issue 4, October 2003
- Miehlabradt, Alexandra Overy; Mary McVay: Developing Commercial Markets for Business Development Services, Seminar Reader for the Third Annual Seminar, Turin, Italy 9-13 September 2003,
- Neubert, Susanne: Akteurszentrierte Wirkungsanalyse in der Entwicklungszusammenarbeit. Manuscript May 2004 a, for DED-Brief, to be published.
- Neubert, Susanne: Akteurszentrierte Wirkungsanalyse in der politischen Entwicklungs- und Verwaltungszusammenarbeit, Manuscript May 2004b, to be published by Nomos-Verlag.
- Neubert, Susanne: Methods for Impact Assessment for Projects and Programms (MAPP), presentation at the annual workshop: Arbeitskreis Entwicklung und Verwaltung Universität Konstanz, 9. – 10. Oktober 2003.
- Neubert, Susanne: Die soziale Wirkungsanalyse in armutsorientierten Projekten. Deutsches Institut für Entwicklungspolitik (DIE/GDI), Bd. 116, Weltforum Verlag, Köln 1999.
- Neubert, Susanne: MAPP: A New Method for Impact Assessment of Poverty Alleviation Projects, German Development Institute (GDI), Berlin mimeo 1999.

Oldsman, Eric: Assessing the Poverty Impact of Small Enterprise Initiatives, prepared for the Working Group for Impact Measurement and Performance, Committee of Donor Agencies for Small Enterprise Development, Nexus Associates, Inc. 25 December 2003, 24p.

Pogge, Thomas W.: Ein Dollar am Tag. Von den Schwierigkeiten, die Weltarmut zu berechnen, in: Neue Zürcher Zeitung, 4./5. Januar 2003, p. 65.

Ranis, Gustav: Human Development and Economic Growth. New Haven: Yale University, Economic Growth Centre, Center Discussion Paper No. 887, May 2004.

Sen, Amartya: Development as freedom. New York: Knopf 1999. (dt. Ökonomie für den Menschen. Wege zur Gerechtigkeit und Solidarität in der Marktwirtschaft, München: Hanser 2000, dtv 2002.

SDC: Guidelines Evaluation and Controlling: Programme Cycle Management, PMET, Controlling and Independent Evaluation.

SDC: English Glossary: 27 most relevant terms related to evaluation & controlling in SDC.

SDC: Key questions and indicators, series of working instruments for planning, evaluation, monitoring, transference into action, March 1999.

Simonowitz, Anton: Virtual Meeting on Impact Assessments Methodologies, background paper, mimeo 21p. (download under www.ids.ac.uk/impact/indexguide.html)

Sinning, Mathias: Kausalanalytische Ansätze zur Erfolgsermittlung armutsrelevanter Massnahmen, Improving Capacities for Poverty Research (IMCAP), University of Colombo, Sri Lanka, on behalf of GTZ, at Heidelberg University, May 2003 (www.wiram.de/wirkung/wirkung.htm).

The Social Impact of Microfinance, Special Issue, Journal of International Development, Vol. 16, Issue 3, April 2004.

Späth, Brigitte: Klein- und Mittelindustrieförderung in Entwicklungsländern. Entstehung und Umsetzung multilateraler Politikansätze durch die UNIDO und die ILO. München: Weltforumverlag, ifo Studien zur Entwicklungsforschung Nr. 32, 1997.

Vahlhaus, Martina: Monitoring impacts – a participatory approach, in: brainstorm, CEFÉ-magazine, Issue 1, April 1999, pp. 4-7.

Vahlhaus, Martina; Thomas Kuby: Orientierungsrahmen für das Wirkungsmonitoring in Projekten der Wirtschafts- und Beschäftigungsförderung unter besonderer Berücksichtigung armutsmindernder Wirkungen, Hrsg. Müller-Glodde, Petra, GTZ, 2000

Warr, Peter G.: Is growth good for the poor? Thailand's boom and bust, in: International Journal of Social economics, Vol. 27, Nr. 7/8/9/10, 2000, pp. 863-877.

Woller, Gary and Robert Parsons: Assessing the Community Economic Impact of Microfinance Institutions, in: Journal of Developmental Entrepreneurship, Vol. 2 – 2002 – No. 2, pp. 133-150.

World Bank 2003: Toward Country-led Development. A multi-partner Evaluation of the Comprehensive Development Framework, Synthesis Report, Washington D.C. World Bank.

World Bank 2002: Monitoring & Evaluation. Some Tools, Methods & Approaches, Washington D.C. World Bank.

Wright, Kate and James Copestake: Impact assessment of microfinance using qualitative data: communicating between social scientists and practitioners using the QUIP, in: Journal of International Development, Vol. 16 2004 Issue 3, April, pp. 355-367.