

Herausgeber:

Prof. Dr. Dr. h. c. Dieter Oberndörfer

Prof. Dr. Theodor Hanf

Prof. Dr. Heribert Weiland

Arbeitskreis „Evaluation von Entwicklungs-
politik“ DeGEval – Deutsche Gesellschaft
für Evaluation (Hrsg.)

**Verfahren der Wirkungsanalyse.
Ein Handbuch für die
entwicklungspolitische Praxis**

Die Publikation dieses Handbuchs wurde finanziell gefördert von:
Evangelischer Entwicklungsdienst (EED); Gesellschaft für Technische
Zusammenarbeit (GTZ); Kreditanstalt für Wiederaufbau (KfW); Misereor.



Redaktionsteam

Christian Berg, Bernward Causemann, Dr. Dirk Guenther, Thomas
Hochgesang, Regina Müller, Dr. Susanne Neubert und Martin Quack

Bibliografische Information der Deutschen Bibliothek:

Die Deutsche Bibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte
bibliografische Daten sind im Internet über
<http://dnb.d-nb.de> abrufbar.

ISBN 978-3-928597-54-8

© 2010 Arnold-Bergstraesser-Institut, Freiburg i. Br.

Alle Rechte vorbehalten

Druck: E & B printware, Digital- und Schnelldruck Gesellschaft mbH
Käppelestraße 10, 76131 Karlsruhe

Vorwort

Mit dieser Publikation dokumentiert die Arbeitsgruppe „Wirkungsanalyse“ des Arbeitskreises „Evaluation von Entwicklungspolitik“ der DeGEval die Ergebnisse ihrer dreijährigen Beschäftigung mit konkreten Wirkungsanalyseverfahren und -designs. Ausgangspunkt und Motivation der Arbeitsgruppe war das Bedürfnis nach einer systematischen Übersicht über erprobte Verfahren – einem Handbuch für die entwicklungspolitische Evaluations- und Managementpraxis. Die Arbeitsgruppe bestand aus Vertreter/-innen der staatlichen und nicht-staatlichen Entwicklungszusammenarbeit (EZ) und der Wissenschaft sowie aus freien Gutachter/-innen und Berater/-innen.

Die anvisierte Leserschaft dieses Handbuchs ist die entwicklungspolitische Fachwelt, die mit der Analyse von tatsächlichen (de facto) empirisch nachprüfbaren Wirkungen befasst ist und die einen schnellen, informativen Zugriff auf das Vorhandene benötigt. Dies sind in erster Linie Projekt- und Programmverantwortliche in den Entwicklungsorganisationen in Geber- und Empfängerländern, die gehalten sind, die Wirkungen von Entwicklungsmaßnahmen nachzuweisen und in ihren Berichten zu dokumentieren; Mitarbeiter/-innen von Evaluierungs- und Qualitätsmanagementeinheiten, die Wirkungsanalysen zumeist als Teil von (Querschnitts-)Evaluationen in Auftrag geben; Wissenschaftler/-innen, die bei dem Design von empirischen Wirkungsstudien beraten oder diese selbst durchführen; sowie Gutachter/-innen und Berater/-innen, die vor der Aufgabe stehen, Wirkungsanalyseverfahren anzuwenden oder in einer Organisation einzuführen.

In dem vorliegenden Handbuch wurde versucht, alle, insbesondere im deutschsprachigen Kontext bekannten und in der Praxis erprobten, gut beschriebenen Wirkungsanalyseverfahren aufzunehmen, wobei ein Anspruch auf Vollständigkeit nicht erhoben werden kann. Dabei liegt der Schwerpunkt nicht nur auf qualitativen und partizipativen Verfahren, die selbst oft zusätzliche quantitative Elemente enthalten, sondern es werden auch quantitative experimentelle und nicht-experimentelle Designs beschrieben, die mitunter als „rigorose“ Verfahren bezeichnet werden. Dabei werden hier keine Bewertungen im Sinne von gut oder schlecht ausgesprochen. Ziel ist es stattdessen, den Nutzern dieses Handbuchs die Möglichkeit zu geben, Verfahren nach rein sachlichen und pragmatischen Erwägungen für einen bestimmten Untersuchungszweck oder -auftrag auszuwählen.

Es werden bewusst Verfahren ausgeklammert, die ausschließlich der Vorab- (ex ante) Abschätzung von Wirkungen in der Planungsphase von Entwicklungsmaßnahmen dienen. Dabei können die Verfahren Bestandteil von Monitoringsystemen (z. B. auch spezifischer Organisationen) sein oder

im Rahmen von Evaluationen oder Wirkungsstudien punktuell zum Einsatz kommen.

Das Handbuch gliedert sich in drei Teile. Im ersten Abschnitt werden die derzeitigen politischen, methodischen und praktischen Herausforderungen für Wirkungsanalysen kurz skizziert. Im zentralen zweiten Abschnitt werden verschiedene Verfahren der Wirkungsanalyse überwiegend in Form von einheitlich strukturierten ‚Steckbriefen‘ dargestellt. Jeder Steckbrief enthält Hinweise auf weiterführende Literatur und sonstige Materialien. Im abschließenden Teil werden die wesentlichen Gemeinsamkeiten und Unterschiede der Verfahren zusammengefasst. Am Ende findet sich eine kurze Schlussfolgerung und ein Glossar mit den wichtigsten Begriffen.

Die Arbeitsgruppe dankt dem Redaktionsteam – Christian Berg (freier Gutachter), Bernward Causemann (freier Gutachter), Dr. Dirk Guenther (Welthungerhilfe), Thomas Hochgesang (CBM), Regina Müller (freie Gutachterin), Dr. Susanne Neubert (DIE) und Martin Quack (Universität Köln) – für die Verfassung der Kapitel 1 und 3 sowie die Zusammenführung der mündlichen und schriftlichen Beiträge zu diesem Handbuch. Eine Liste sämtlicher Teilnehmer/-innen der Arbeitsgruppe sowie Informationen zu den Autor/-innen finden sich im Anhang.

Dem Arnold-Bergstraesser-Institut wird für den Satz des Manuskripts und die Unterstützung bei der Veröffentlichung gedankt.

Inhaltsverzeichnis

Verzeichnis der Abbildungen	10
Verzeichnis der Abkürzungen	11
1. Herausforderungen für Wirkungsanalysen	15
1.1 Politische Herausforderungen	16
Millennium Declaration und Aid-Effectiveness-Debatte	16
Veränderungen in der EZ-Architektur	17
Internationale Vereinheitlichung von Standards und Kriterien	19
Forderung des Bundesrechnungshofes nach Erfolgskontrolle	20
1.2 Methodische Herausforderungen	20
Standards und Prinzipien	21
Das Zuordnungsproblem	22
Nachweis- und Vergleichsmöglichkeiten	23
1.3 Praktische Herausforderungen	26
Wirkungserfassung im Rahmen von Monitoring und Evaluierungen	26
Wahl eines geeigneten Verfahrens der Wirkungsanalyse	27
Praktische Herausforderungen für die Evaluierer/-innen	29
Institutionelle Verankerung der Wirkungsanalyse	31
Mittelaufwand	32
2. Wirkungsanalysen in der Praxis	33
2.1 Struktur der Verfahrens-Steckbriefe	33
2.2 Verfahren der Wirkungsanalyse	35
CEval-Ansatz zur Wirkungsevaluation / Stockmann'scher Ansatz	36
Do No Harm (DNH)	50
e-VAL [®]	56
Impact Monitoring and Assessment (IMA) nach Herweg/Steiner	66
Impact Pathway	72
KAP-Studien	79
Method for Impact Assessment of Programmes and Projects (MAPP)	88

Most Significant Change (MSC)	97
NGO-IDEAS Impact Toolbox	105
Outcome Mapping (OM)	111
Participatory Impact Monitoring (PIM) / FAKT	120
Participatory Impact Monitoring (PIM) / SLE	128
Participatory Livelihood System Analysis (PaLSA)	136
Participatory Methodology for Measuring the Contribution of Volunteering to Development (MCVD)	142
Peace and Conflict Assessment (PCA)	149
Programm-integriertes Planungs-, Monitoring- und Evaluierungssystem (PriME) von InWEnt	155
Quantitative experimentelle Designs	162
Quantitative nicht-experimentelle Designs	167
Real Time Evaluation (RTE)	180
Verbleibsstudien (Tracer Studies)	187
Wirkungsorientiertes Monitoring (WoM) der GTZ	192
2.3 Weitere Ansätze der Wirkungsanalyse	200
Wirkungsanalyse von EZ in Konfliktsituationen	200
Wirkungsanalyse von Ernährungssicherungsmaßnahmen	202
Erfassung der Wirkungen von (Multi-)Sektorprogrammen	203
Methodisches Vorgehen bei der Ermittlung von Politikeinfluss	205
Poverty and Social Impact Analysis (PSIA) und Poverty Impact Assessment (PIA)	211
3. Synthese – Verfahren der Wirkungsanalyse in der Zusammenschau	215
3.1 Gemeinsamkeiten und Unterschiede in den Verfahren	215
Grad der Konkretion und institutionellen Verankerung	215
Zugänglichkeit und Voraussetzungen	216
Anwendbarkeit für bestimmte Sektoren, Maßnahmenebenen und -typen	217
Eignung für die Erfassung verschiedener Arten von Wirkung	219
Ausgangspunkt der Analyse und Begründung der Wirksamkeit	221
Verwendung quantitativer und qualitativer Methoden	222
Art der Validierung von Wirkungsdaten/-informationen	223
Möglichkeit der Aggregation von Wirkungsdaten/-informationen	224

Eignung für die Erfassung von Zuständen und Prozessen	224
Grad der Beteiligung von Trägern und Zielgruppen an der Wirkungsanalyse	225
Verknüpfungen und Weiterentwicklungen von Verfahren	226
3.2 Anwendungs- und Verbesserungsmöglichkeiten	226
3.3 Überblick über Anwendungsmöglichkeiten und Voraussetzungen der Verfahren	228
Literatur	234
Glossar	240
Mitglieder der AG „Wirkungsanalyse“	244
Autorinnen und Autoren	245

Verzeichnis der Abbildungen

Abbildung 1:	Darstellung der partizipativen Einbindung der Evaluierten in den Evaluationsprozess gemäß des Stockmann'schen Evaluationsansatzes	42
Abbildung 2:	Übersicht über Do No Harm	53
Abbildung 3:	Beispiel für die Datenauswertung bei e-VAL®	61
Abbildung 4:	Beispiel für ein Spinnendiagramm in IMA nach Herweg/Steiner	68
Abbildung 5:	Impact Pathway am vereinfachten Beispiel der Entwicklung von Privatwirtschaft in Mosambik im Rahmen des DFID Business Linkages Challenge Fund	75
Abbildung 6:	Beispiel einer Einflussmatrix bei MAPP	92
Abbildung 7:	Beispiel eines Entwicklungs- und Wirkungsprofils bei MAPP	93
Abbildung 8:	Die MSC-Technik im Überblick	102
Abbildung 9:	Interne Kontrollmechanismen zur Validierung von Ergebnissen bei NGO-IDEAS	108
Abbildung 10:	Schritte zur Einführung und Durchführung von PIM / FAKT	122
Abbildung 11:	Verfahrensschritte bei PIM / SLE	131
Abbildung 12:	Durchführungsphasen von PaLSA	139
Abbildung 13:	Workshopebenen und -aktivitäten bei MCVD	145
Abbildung 14:	Aid for Peace Framework	151
Abbildung 15:	PM&E-Zyklus und Arbeitsschritte bei PriME	158
Abbildung 16:	Verbleibsstudien integriert in MIS von Berufsbildungseinrichtungen	189
Abbildung 17:	Wirkungsmodell der GTZ	194
Abbildung 18:	7 Schritte eines wirkungsorientierten Monitoring in der GTZ	197
Abbildung 19:	Relevante Wirkungsdimensionen erster und zweiter Ordnung bei Vorhaben mit zwei und mehreren Partnern	207
Abbildung 20:	Raster zur Messung von Politikeinfluss	209
Abbildung 21:	Poverty Impact Assessment (PIA) Module	213

Verzeichnis der Abkürzungen

3ie	International Initiative for Impact Evaluation
ABI	Arnold-Bergstraesser-Institut
ADRA	Adventist Development and Relief Agency
AFC	AFC Consultants International
AfP	Aid for Peace
AGEG	Arbeitsgemeinschaft entwicklungspolitischer Gutachter
AGEH	Arbeitsgemeinschaft für Entwicklungshilfe
AGKED	Arbeitsgemeinschaft kirchlicher Entwicklungsdienst
AI	Appreciative Inquiry
ALRMP	Arid Lands Resource Management Project
AURA	Auftragsrahmen
AVI	Australian Volunteers International
BfdW	Brot für die Welt
BGR	Bundesanstalt für Geowissenschaften und Rohstoffe
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
BOND	British Overseas NGOs for Development
CARE	Cooperative for American Remittances to Europe
CBM	Christoffel Blindenmission – Internationale Organisation für Behinderung und Entwicklung
CBO	Community-Based Organisation
CCI	Canadian Crossroads International
CCIC	Canadian Council for International Cooperation
CDA	Collaborative for Development Action
CDE	Centre for Development and Environment
CEVAL	Centrum für Evaluation
CGIAR	Consultative Group on International Agricultural Research
CONCORD	European NGO Confederation for Relief and Development
DAAD	Deutscher Akademischer Austauschdienst
DAC	Development Assistance Committee
DEC	Disaster Emergency Committee
DED	Deutscher Entwicklungsdienst
DeGEval	Deutsche Gesellschaft für Evaluation
DEZA	Direktion für Entwicklung und Zusammenarbeit
DFID	Department for International Development
DiD	Differenzen in Differenzen
DIE	Deutsches Institut für Entwicklungspolitik
DNH	Do No Harm
DO	Durchführungsorganisation

DSE	Deutsche Stiftung für Entwicklung
EASY ECO	Evaluation of Sustainability – European Conferences and Training Courses
EDV	Elektronische Datenverarbeitung
EED	Evangelischer Entwicklungsdienst
EIMS	Evaluation Itérative avec Mini-Séminaires
EINS	Eins Entwicklungspolitik
ENÜH	Not- und Übergangshilfe
EON	Entwicklungsorientierte Nothilfe
EPAU	Evaluation and Policy Analysis Unit
EQM	Evaluierung und Qualitätsmanagement
e-VAL®	Computergestütztes Interviewverfahren der GTZ
EZ	Entwicklungszusammenarbeit
EZE	Evangelische Zentralstelle für Entwicklungshilfe
FAKT	FAKT Consult for Management, Training and Technologies
FAO	Food and Agricultural Organization
FAQ	Frequently asked questions
FES	Friedrich-Ebert-Stiftung
FEZ	Food Economy Zones
FriEnt	Gruppe Friedensentwicklung
FZ	Finanzielle Zusammenarbeit
GATE	German Appropriate Technology and Ecoefficiency Programme
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit
HBS	Heinrich-Böll-Stiftung
HEA	Household Economy Approach
HFES	Household Food Economy Survey
IDEAs	Impact on Development, Empowerment and Actions
IDD	International Development Department der Universität Birmingham
IDRC	International Development Research Centre
IEM	Implicit Ethical Message
IFAD	International Fund for Agricultural Development
ILO	International Labour Organization
IMA	Impact Monitoring and Assessment
INTRAC	International NGO Training and Research Centre
InWent	Internationale Weiterbildung und Entwicklung
IWRM	International Water and Resources Management
IZA	Institut zur Zukunft der Arbeit
JOCV	Japan Overseas Co-operation Volunteers
KAP (KAB)	knowledge, attitude, practice (behaviour)

KfW	Kreditanstalt für Wiederaufbau
LFA	Logical Framework Analysis
MAPP	Method for Impact Assessment of Programmes and Projects
MCVD	Participatory Methodology for Measuring the Contribution of Volunteering to Development
MDG	Millennium Development Goal
M&E	Monitoring- und Evaluierungsverfahren
MS	Multi-Sectoral
MSC	Most Significant Change
MYRADA	Mysore Resettlement and Development Agency
NGO	Non-Government Organization
NONIE	Network of Networks on Impact Evaluation
NRO	Nichtregierungsorganisation
NYSASDRI	National Youth Service Action and Social Development Research Institute
NZARE	New Zealand Association for Research in Education
ODI	Overseas Development Institute
OECD	Organization for Economic Cooperation and Development
OM	Outcome Mapping
OLS	Ordinary Least Squares
OXFAM	Oxford Committee for Famine Relief
PAG	Performance Appraisal of the Groups
PaLSA	Participatory Livelihood System Analysis
PANgo	Performance Assessment of NGOs
PBA	Programme Based Approaches
PCA	Peace and Conflict Assessment
PCIA	Peace and Conflict Impact Assessment
PCM	Project Cycle Management
PGF	Programmorientierte Gemeinschaftsfinanzierung
PIA	Poverty Impact Assessment
PIM	Participatory Impact Monitoring
PM&E	Planung, Monitoring und Evaluierung
PRA	Participatory Rural Appraisal
PrIME	Programm-integriertes Planungs-, Monitoring- und Evaluierungssystem
PRSP	Poverty Reduction Strategy Paper
PSIA	Poverty and Social Impact Analysis
PSM	Propensitätsscorematching
PTB	Physikalisch Technische Bundesanstalt
PWR	Participatory Wealth Ranking
RAN	Rapid Assessment of Nutrition

RBM	Results Based Management
RDD	Regressionsdiskontinuitätsansatz
RNA	Rapid Nutrition Assessment
RNRRS	Renewable Natural Resources Research Strategy
RPA	Rapid Portfolio Assessment
RRA	Rapid Rural Appraisal
RSA	Rapid Sectoral and Rapid Portfolio Assessment
RT	Ressourcentransfer
RTE	Real Time Evaluation
SAGE	Situational Analysis and Goal Establishment
SCRIA	Supply Chain Relationships in Action
Seco	State Secretariat for Economic Affairs
SHG	Selbsthilfegruppe
SHO	Self-help Organization
SIA	Social Impact Assessment
SIDA	Swedish International Development Cooperation Agency
SLE	Seminar für Ländliche Entwicklung
SMART	Specific, Measurable, Attainable, Relevant, Time bound
SRHR	Sexuelle, reproduktive Gesundheit und Rechte
STIPA	Support for Tropical Initiatives in Poverty Alleviation
SWap	Sector-wide Approach
SWAP	Soziale Wirkungsanalyse in armutsorientierten Projekten
SWOT	Strengths, Weaknesses, Opportunities, Threats
TOR	Terms of Reference
TZ	Technische Zusammenarbeit
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations International Children's Emergency Fund
UNV	United Nations Volunteers
URD	Urgence Réhabilitation Développement
VENRO	Verband Entwicklungspolitik deutscher Nichtregierungsorganisationen
VN	Vereinte Nationen
VSA	Volunteer Service Abroad
WCD	World Commission on Dams
WFP	World Food Programme
WoM	Wirkungsorientiertes Monitoring
ZOPP	Zielorientierte Projektplanung

1. Herausforderungen für Wirkungsanalysen

Die Beschäftigung mit Ansätzen, Verfahren und Methoden der Wirkungsanalyse in der Entwicklungszusammenarbeit (EZ) ist nicht neu, und dennoch hat die Debatte um Wirkungsorientierung und Wirkungsnachweis in den letzten Jahren an Dynamik zugenommen.

Aus **politischer Perspektive** hat nicht nur die Pflicht zur Rechenschaftslegung gegenüber Gebern (z. B. Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ)), dem Parlament und der Öffentlichkeit weiter zugenommen. Auch internationale Vereinbarungen zur verbesserten Wirkungsorientierung von Entwicklungsmaßnahmen und die hochrangigen Konferenzen zur Wirksamkeit der EZ, zuletzt in Accra 2008, halten die Akteure der EZ dazu an, diesem Anspruch zu entsprechen. Veränderungen in der EZ-Architektur sowie die zunehmenden Bestrebungen nach Vereinheitlichung von Evaluierungskriterien und die Debatte um Evaluierungsstandards erfordern ebenfalls eine weitergehende Orientierung auf Wirkung und eine systematische Herangehensweise bei deren Erfassung und Dokumentation (siehe Abschnitt 1.1). Dabei haben staatliche, nicht-staatliche oder internationale Organisationen, die Entwicklungsmaßnahmen mit eigenem Personal vor Ort durchführen (Durchführungsorganisationen i. w. S.), direktere Möglichkeiten für die Um- bzw. Durchsetzung der Wirkungsorientierung als Organisationen, die Partnerorganisationen fördern (Förderorganisationen).

Aus **methodischer Perspektive**, die in diesem Handbuch im Vordergrund steht (siehe Abschnitt 1.2), besteht nach wie vor die Herausforderung darin, Wirkungen von Entwicklungsmaßnahmen hinreichend zuverlässig und gültig zu belegen¹, und aus der **Perspektive der Praxis** (siehe Abschnitt 1.3) gilt es, die methodischen Anforderungen mit den realen Voraussetzungen in Einklang zu bringen sowie die notwendigen organisatorischen und finanziellen Voraussetzungen für gut fundierte Wirkungsanalysen – integriert in Planungs-, Monitoring- und Evaluierungssysteme wie auch als allein stehende Studien – zu schaffen.

¹ Aktuell wird – seit Erscheinen der Publikation „When will we ever learn?“ des Center for Global Development (2006) – das Thema „Nachweis von Wirkungen“ unter dem Stichwort „(Rigorous) Impact Evaluation“ in der internationalen und deutschen EZ wieder verstärkt diskutiert.

1.1 Politische Herausforderungen

Millennium Declaration und Aid-Effectiveness-Debatte

Mit der im September 2000 verabschiedeten Millenniumserklärung und den 2001 formulierten und mit Indikatoren bestückten Entwicklungszielen bis zum Jahr 2015, den Millennium Development Goals (MDGs), hat die Ausrichtung der Entwicklungszusammenarbeit auf Wirkungen eine neue Verbindlichkeit erfahren. In darauf folgenden einschlägigen internationalen Konferenzen, wie dem International Roundtable on Managing for Development Results in Marrakesch (2004), der Konferenz zur Aid Effectiveness in Paris (2005) und dem High Level Forum in Accra (2008), wurde diese Ausrichtung weiter konkretisiert:

- Für die Wirkungsorientierung ist ein kohärenter Ansatz entscheidend, der nicht nur die ex-ante Phase eines Vorhabens umfasst (Planung und Vorab-Einschätzung der Wirkungen, Erstellung von *baselines*), sondern ebenso die Durchführungsphase (wirkungsorientiertes Monitoring anhand von Erfolgszielen und -indikatoren) und die Abschlussphase (Schluss- und ex-post-Einschätzung der Wirkungen). Hierzu müssen Wirkungsbeziehungen wie z. B. Wirkungsketten berücksichtigt und qualitative und quantitative Wirkungsindikatoren realistisch ausgewählt werden.²
- Wirkungsorientierung bedeutet, Management und Durchführung von Entwicklungsmaßnahmen auf gewünschte Wirkungen hin auszurichten und Informationen über Wirkungen für verbesserte Entscheidungsfindung zu nutzen.³

Die Erkenntnis, dass einzelne Projekte und Programme oft als hochwirksam evaluiert werden, positive gesamtwirtschaftliche und -gesellschaftliche Auswirkungen der EZ aber oftmals im hoch aggregierten Ländervergleich nicht nachweisbar sind, bereitet nicht nur der Wissenschaft und der Politik Kopfzerbrechen. Auch die Praxis der EZ wird zunehmend davon beeinflusst. Der aus diesem "Mikro-Makro-Paradoxon"⁴ resultierende, im Zusammenhang mit der Paris-Erklärung deutlich erhöhte Druck auf die deutsche EZ, den Wirkungsnachweis zu verbessern und vollständiger zu machen, betrifft nicht nur die staatliche Entwicklungszusammenarbeit (hier das BMZ und dessen Durchführungsorganisationen⁵, sondern auch deutsche Nichtre-

² Siehe Joint Marrakech Memorandum (2004).

³ Siehe Paris Declaration on Aid Effectiveness (OECD 2005), Ziffer 43.

⁴ Siehe z. B. Wolff, Jürgen H. (2005) und Mosley, P. (1987).

⁵ Formale Grundlage für die Evaluierungsaktivitäten – und damit auch für die Wirkungserfassung und -dokumentation – des BMZ und der staatlichen bilateralen Durchführungsorgani-

gierungsorganisationen (NRO).⁶ 2007 waren mehrere NRO in eine vom BMZ beauftragte „Systemprüfung der Evaluation in der deutschen Entwicklungszusammenarbeit“ einbezogen, unter anderem mit der Frage nach den Strukturen und Verfahren, die der Wirkungsanalyse der von ihnen durchgeführten oder geförderten Entwicklungsmaßnahmen dienen.

Allerdings wurden deutsche zivilgesellschaftliche Organisationen nicht erst durch die Paris-Erklärung zum Thema Wirkungsanalyse aktiv: Schon Ende der 1980er Jahre und verstärkt ab 2000 haben NRO das Thema Wirkungsbeobachtung angesprochen. So haben zum Beispiel kirchliche NRO wohl als erste NRO dieses Thema in ihrem Arbeitsbuch thematisiert.⁷ Seit 2004 entwickelt die NRO-Initiative NGO-IDEAs⁸ gemeinsam mit Partnern aus Entwicklungsländern Konzepte und Instrumente der wirkungsorientierten Projektsteuerung (einschließlich Wirkungsmonitoring) auf partizipativer Basis. Diese Initiative wird durch das BMZ gefördert und von VENRO unterstützt.⁹ Andere NRO, wie zum Beispiel die Welthungerhilfe, folgten.

Veränderungen in der EZ-Architektur

Mit den internationalen Verpflichtungen gingen, zumindest in der staatlichen EZ, eine konzeptionelle Neuausrichtung und Veränderungen in den entwicklungspolitischen Verfahren und Instrumenten einher, die besondere Herausforderungen an Wirkungsanalyse stellen: Programmbildung (Länderprogramme), Sektorprogramme (*sector-wide approaches*), programmorientierte Gemeinschaftsfinanzierung (*program-based approach*) oder auch Budgethilfe. Bislang stehen nur wenige Ansätze und noch keine völlig ausgereiften

sationen (DO: GTZ, KfW, DED, InWEnt, PTB und BGR) sind die FZ-/TZ-Leitlinien in ihrer jeweils aktuellsten Fassung. Darin ist u. a. festgelegt, dass laufende, abschließende und ex-post-Evaluierungen von Projekten und Programmen auf der Basis einer aussagekräftigen Stichprobe valide Aussagen zur entwicklungspolitischen Wirksamkeit und zu positiven und negativen Lernerfahrungen der von den jeweiligen DO durchgeführten Entwicklungsmaßnahmen ermöglichen sollen, die auch als Grundlage für weiter gehende Analysen dienen können.

⁶ Allerdings gehen die zivilgesellschaftlichen Organisationen in Deutschland von einem breiteren Verständnis von Wirksamkeit aus und beziehen die Reichweite der Paris-Erklärung auf die Bereitstellung öffentlicher Entwicklungshilfe (ODA), siehe Post und Warning (2007), VENRO (2007).

⁷ AGKED/Misereor (1992), Dolzer u. a. (1998), Gohl (2000), ICCO (2000).

⁸ NGO-Impact on Development Empowerment and Actions (<http://www.ngo-ideas.net/home>).

⁹ Karl-Kübel-Stiftung (2007), VENRO (2007), Gohl (2007), Santha (2007).

Verfahren zur Analyse der Wirksamkeit solcher Instrumente zur Verfügung.¹⁰

Die aus den Veränderungen resultierenden Ansprüche an Verfahren der Wirkungsanalyse sind dementsprechend größer geworden. Da sich alle, bzw. zumindest staatliche, EZ-Organisationen letztendlich an der Erreichung der MDGs messen lassen müssen oder wollen, wird von den eingesetzten Verfahren erwartet, dass deren Ergebnisse aggregiert und nach unterschiedlichen Gesichtspunkten und auf mehreren Ebenen ausgewertet werden können. Auch steigt mit zunehmenden Budget- und anderen Gemeinschaftsfinanzierungen der Anspruch und die Notwendigkeit, Wirkungsanalysen vermehrt gemeinschaftlich, d. h. mit Hilfe mehrerer Akteure unter Beteiligung auch der Partnerregierungen durchzuführen.¹¹

Es sind noch etliche Fragen offen, die sich aus der neuen EZ-Architektur ergeben: Wie kann sich die Zivilgesellschaft in den Partnerländern z. B. an Sektorvorhaben und dem Monitoring der Wirkungen solcher Instrumente, wie etwa Budgethilfe, beteiligen? Welche gesellschaftspolitischen Wirkungen bringt ein stark auf Harmonisierung der öffentlichen EZ ausgerichtetes Entwicklungsparadigma mit sich und wie lassen sich diese Wirkungen erfassen und dokumentieren oder auch als spezifischer „Beitrag“ bestimmten Akteure zuordnen und wäre dies überhaupt sinnvoll? Wie kann aus vornehmlich auf Regierungshandeln abzielendem *ownership* eine „demokratisch legitimierte Eigenverantwortung“ unter voller Einbeziehung der Zivilgesellschaft werden, und wie können solche Wirkungen gemessen werden? Die zivilgesellschaftlichen Organisationen in Deutschland, deren Partnerorganisationen autonom Projekte in Entwicklungsländern durchführen, haben im Dialog im Rahmen der *aid-effectiveness*-Debatte und EZ-Architektur ihre eigene kritische Position eingebracht.¹² Nicht nur bei zivilgesellschaftlichen Organisationen wächst hierbei die Befürchtung, dass die zunehmend an quantitativen Wirkungsindikatoren ausgerichteten Berichtspflichten so aufwändig werden, dass sie zu viele Ressourcen für die Admi-

¹⁰ So ist beispielsweise noch weitgehend offen, wie die Wirkungen, die diese Änderungen der EZ-Architektur nach sich ziehen, adäquat erfasst werden können. Ob beispielsweise die Partner(regierungen) in gemeinsamen Programmen und in der Budgetfinanzierung mehr *ownership* empfinden bzw. übernehmen als in ‚traditionellen‘ Projekten und Programmen bleibt nachzuweisen. Auch welche Auswirkungen die Verpflichtung zur gegenseitigen Rechenschaftslegung auf den Einfluss von Bürgern und Parlamenten auf die Allokation von Entwicklungsgeldern in den Partnerländern haben wird, sollte mittels geeigneter Verfahren analysiert werden können.

¹¹ Siehe Zintl (2006).

¹² Siehe The North-South Civil Society Dialogue on Aid Effectiveness (2007) und The Advisory Group on Civil Society & Aid Effectiveness (2008).

nistration statt für die Zielgruppen verbrauchen sowie die notwendige Flexibilität in der partizipativen Entwicklungsarbeit behindern könnten.¹³

Internationale Vereinheitlichung von Standards und Kriterien

In den vergangenen Jahren haben internationale wie nationale Anstrengungen zu einer stärkeren Vereinheitlichung von Standards und Kriterien für das Monitoring und die Evaluierung von Wirkungen der EZ¹⁴ geführt, die Orientierung für Verfahren der Wirkungsanalyse bieten:

Results Based Management (RBM), das sich stark am *logframe approach* orientiert, beinhaltet eine möglichst objektive und für das Management nützliche Messung der Zielerreichung (*outputs, outcome und impact*) von Entwicklungsmaßnahmen anhand SMARTer Indikatoren.¹⁵

Die Qualitätsstandards des DAC (Development Assistance Committee) für Evaluierung¹⁶, die weitgehend mit den DeGEval-Standards für Evaluierung¹⁷ übereinstimmen, legen u. a. nahe

- eine Analyse des (sozio-)politischen und institutionellen Kontextes und anderer Einflüsse auf die Wirkungen der Entwicklungsmaßnahme,
- eine detaillierte Beschreibung der Methoden und Techniken (einschließlich der Stichprobenziehung) für die Sammlung und Verarbeitung von (Wirkungs-)Daten und -Informationen,
- die Beteiligung aller wichtigen Akteure an der (Wirkungs-)Untersuchung,
- Transparenz der Informationsquellen sowie
- die Validierung von Informationsquellen und Daten.¹⁸

Das BMZ sieht die fünf Evaluierungskriterien des DAC (Relevanz, Effektivität, Effizienz, entwicklungspolitische Wirkung und Nachhaltigkeit) sowie die „3 K“ (Kohärenz, Komplementarität und Koordination) für alle entwick-

¹³ Earle (2004).

¹⁴ Results-Based Management / RBM, DAC Evaluation Quality Standards und Criteria for Evaluating Development Assistance, BMZ Evaluierungskriterien für die deutsche bilaterale Zusammenarbeit.

¹⁵ Siehe z. B. DAC Working Party on Aid Evaluation (2001), S. 6; UN Joint Inspection Unit (2004), S. 14; DAC (2002).

¹⁶ Siehe DAC Evaluation Network (2006).

¹⁷ Siehe DeGEval (2002).

¹⁸ Die DeGEval-Standards betonen zudem besonders die Nützlichkeit (Verständlichkeit, Nachvollziehbarkeit, Nutzung und Nutzen) der Ergebnisse von Evaluierungen und damit auch von Wirkungsanalysen (siehe DeGEval 2002).

lungspolitischen Maßnahmen deutscher Durchführungsorganisationen als verbindlich und damit die Analyse von direkten und indirekten Wirkungen als zentral an.¹⁹

Die Evaluierungseinheiten der Vereinten Nationen (VN), der OECD, der Entwicklungsbanken sowie nationale Evaluierungsgesellschaften arbeiten im Network of Networks on Impact Evaluation (NONIE) an einem gemeinsamen Verständnis und einer Vereinheitlichung der Verfahren der Wirkungsevaluierung. Mit ähnlichen Zielen wurde ebenfalls 2007 die „International Initiative for Impact Evaluation“ (3ie) gegründet, der Regierungen aus Entwicklungsländern ebenso angehören wie staatliche und multilaterale Geber, internationale zivilgesellschaftliche Organisationen und Forschungsinstitutionen. In beiden Netzwerken wurde ursprünglich der Schwerpunkt auf rigorose Verfahren gelegt. Inzwischen entwickelt sich die Diskussion jedoch in die Richtung, dass ein wohl durchdachter Methodemix am ehesten die Liste der Ansprüche an Wirkungsevaluierung erfüllt.

In internationalen Netzwerken unter Beteiligung von CONCORD, InterAction (USA), BOND (Großbritannien) oder CCIC (Kanada) wie auch auf nationaler Ebene, so z. B. bei der VENRO-Transparenzinitiative oder in NGO-IDEAs, arbeiten derzeit deutsche Organisationen der Zivilgesellschaft an Verhaltenskodices und Verfahren mit, die eine Orientierung auf Wirkung festschreiben wollen.

Forderung des Bundesrechnungshofes nach Erfolgskontrolle

Zusätzlich drängt der Bundesrechnungshof seit einigen Jahren verstärkt auf die Beachtung einschlägiger Verwaltungsvorschriften in der Bundeshaushaltsordnung zur Erfolgskontrolle finanzwirksamer Maßnahmen des Bundes. Dabei sind Bundesbehörden wie z. B. das BMZ gehalten, die Zielerreichung, Wirksamkeit und Wirtschaftlichkeit der von ihnen geförderten Maßnahmen begleitend und abschließend zu überprüfen, im Regelfall durch Schreibtischprüfung der Berichte der öffentlichen wie zivilgesellschaftlichen Empfängerorganisationen. Für diese – und deren Partnerorganisationen – ergibt sich daraus die Herausforderung, die entsprechenden Informationen zu generieren und zu dokumentieren.

1.2 Methodische Herausforderungen

Es gibt eine Reihe methodischer Herausforderungen, vor denen Wirkungsanalysen stehen. Der Begriff der „Wirkung“ (im Sinne eines veränderten oder erhaltenen Zustands) beinhaltet bereits, dass es Ursachen gibt, die zu

¹⁹ Siehe BMZ (2006).

ihr führen, und dass sie umgekehrt, d. h. ohne die Ursachen, nicht eingetreten wäre. Nach dieser kontrafaktischen (hypothetischen) Betrachtungsweise kann jede Maßnahme in der Entwicklungszusammenarbeit unterschiedlichste Wirkungen haben, die in ihrer Gesamtheit und als sich gegenseitig beeinflussendes Wirkungsgeflecht in der Praxis nicht vollständig erfasst und beschrieben werden können. Die dafür notwendigen Laborbedingungen sind praktisch nicht erfüllbar und die Konzepte zur Beschreibung der Zusammenhänge nicht vollständig nachvollziehbar.²⁰ Eine Wirkungsanalyse muss deshalb die zu untersuchenden Wirkungen (erwartete positive und befürchtete negative) definieren und darstellen, wie genau sie nachgewiesen werden. Sie sollte zugleich offen für nicht erwartete Wirkungen sein. Ob eine solche detaillierte Wirkungsbeschreibung auf einer im Vorfeld einer Entwicklungsmaßnahme konstruierten Wirkungskette (*program theory*, *theory of change*) beruhen sollte oder ob es besser ist, zu einem späteren Zeitpunkt von den beobachtbaren Veränderungen (oder Zuständen) vor Ort auszugehen und auf dieser Grundlage die Zuordnung zu möglichen Ursachen vorgenommen werden sollte, wird in der Fachwelt kontrovers diskutiert und gehandhabt.²¹

Standards und Prinzipien

Das Evaluierungskriterium „Impact“ von OECD/DAC meint die positiven und negativen, beabsichtigten und unbeabsichtigten entwicklungspolitischen Wirkungen einer Entwicklungsmaßnahme, insbesondere bei Zielgruppen. Dabei sollen auch konkurrierende Einflüsse durch externe Faktoren überprüft werden. Die Genauigkeitsstandards für Evaluation der DeGEval²² sollen (wie vergleichbare internationale Standards auch) gültige, d. h. möglichst aussagekräftige Informationen und Ergebnisse sicherstellen. Für Wirkungsanalysen sind diese Standards besonders wichtig, insbesondere die Anforderung, dass Informationen gültig und zuverlässig sein müssen (Standard G5). Dieser Standard beruht auf Gütekriterien empirischer Sozialforschung: In der quantitativen Forschung sollen Messungen objektiv, zuverlässig (reliabel) und gültig (valide) sein, wobei diese Kriterien nur selten vollständig, sondern nur zu einem bestimmten Grad erfüllt werden können.²³ Die Einhaltung dieser drei Kriterien „Objektivität“, „Reliabilität“ und

²⁰ Siehe Stockmann (2006), S. 104. Zudem würden dann andere wichtige Forderungen, z. B. nach Partizipation der Partner und Zielgruppen sowie Transparenz und Nachvollziehbarkeit für die (Fach)-Öffentlichkeit zurücktreten.

²¹ Siehe hierzu z. B. Neubert, (2004a, S. 2; 2004b).

²² Siehe DeGEval (2002), S. 31-36.

²³ Vgl. Diekmann (2005), S. 216-227.

„Validität“ ist für die Analyse von Wirkungen dann besonders schwierig, wenn die zu evaluierenden Entwicklungsmaßnahmen und deren Ziele komplex sind und die Grundlagen (z. B. Datenbasis) für quantitative Untersuchungen nicht vorliegen. Selbst wenn die Gütekriterien quantitativer Forschung nicht in vollem Ausmaß erfüllt werden können, bedeutet dies jedoch noch keine Beliebigkeit. Kriterien der qualitativen Analyse wie Sättigungsgrad, Triangulation, Schlüssigkeit und intersubjektive Nachvollziehbarkeit des Vorgehens bleiben wichtige Kriterien für Wirkungsanalysen, an denen sie sich messen lassen müssen.²⁴

Das Zuordnungsproblem

Das zentrale Problem bei Wirkungsanalysen ist der Nachweis, dass bestimmte Veränderungen – oder auch „konservierende Wirkungen“, d. h. Wirkungen, die im Erhalt eines Ausgangszustands bestehen²⁵ – kausal mit der zu untersuchenden Entwicklungsmaßnahme zusammenhängen. Selbst wenn festgestellt werden kann, dass die angestrebten Veränderungen eingetreten sind, können andere Faktoren dafür verantwortlich sein. Dabei können mehrere konkurrierende Ursachen unterschieden werden²⁶: endogener Wandel (innerhalb des Untersuchungsgegenstandes, z. B. Gesundung), exogener Wandel (außerhalb des Untersuchungsgegenstandes, z. B. bessere Konjunktur) oder ein bestimmtes Ereignis (z. B. Straßenbau).²⁷

Besonders schwierig kann der Wirkungsnachweis unbeabsichtigter Wirkungen sein. Kompliziert ist auch die Erfassung langfristiger Wirkungen und solcher auf gesellschaftlichen oder strukturellen Ebenen, die sich von der Interventionsebene unterscheiden. „Diffuse“ Wirkungen von Entwicklungsmaßnahmen, für die etwa keine klar abzugrenzenden Zielgruppen definiert werden können, sind ebenfalls schwierig zu messen. Hierunter fallen z. B. Maßnahmen zur Förderung des Investitionsklimas, große Infrastrukturvorhaben wie der Bau eines Flughafens oder von Überlandstraßen

²⁴ In der qualitativen Sozialforschung wurden Vorschläge für veränderte Gütekriterien entwickelt, besonders wichtig ist eine detaillierte Dokumentation des gesamten Evaluationsprozesses (vgl. Appelsmeyer, Kochinka und Straub 1997, S. 734; Neubert 2007; Patton 2002, S. 544-545).

²⁵ Die Aufrechterhaltung von etwas Positivem als wichtige Wirkung eines Vorhabens wird besonders deutlich im Naturschutz und in der Friedensförderung, kann aber in allen Politikbereichen eine Rolle spielen.

²⁶ Siehe Stockmann (2006), S. 106.

²⁷ Eine gemeinsame Publikation von 13 evangelischen Werken aus Nord und Süd (ICCO 2000) empfiehlt daher (S. 51) Sektor- oder Regionalstudien, die von mehreren Organisationen gemeinsam durchgeführt werden.

sowie breit angelegte *Advocacy*- oder politische Arbeit, z. B. zur Demokratisierung.

All diese Herausforderungen werden unter dem Stichwort „Zuordnungsproblem“ diskutiert.²⁸ Hinzu kommen praktische Probleme wie eine mangelhafte Datenlage für die Bestimmung der Ausgangssituation oder Probleme bei der Operationalisierung von Zielkonzepten (z. B. Messung von Armutsminderung, Demokratisierung o. ä.) und Maßnahmen.

Nachweis- und Vergleichsmöglichkeiten

Da die tatsächliche Entwicklung (mit entwicklungspolitischer Maßnahme) zumeist nicht direkt mit der kontrafaktischen Entwicklung (ohne Maßnahme) verglichen werden kann, werden bei Wirkungsanalysen verschiedene Alternativen zum Wirkungsnachweis verwendet. Häufig wird dabei die Ist-Situation einer Vergleichssituation gegenübergestellt. Dies ist prinzipiell über einen „mit-ohne“-Vergleich oder einen „vorher-nachher“-Vergleich möglich.

Dabei können die Verfahren unterschiedliche Ausschnitte der Wirkungen (geplante/ungeplante, direkte/indirekte, positive/negative) berücksichtigen. Es gibt zwei mögliche Richtungen des Nachweises: Entweder wird die Wirkungskette von der Entwicklungsmaßnahme her untersucht, vom Output über Outcomes bis zu möglichen *Impacts*, oft anhand von Wirkungshypothesen (eher deduktiv). Oder es wird zunächst untersucht, wie sich die Situation, zu deren Verbesserung die Maßnahme durchgeführt wurde, entwickelt hat (z. B. der Gesundheitszustand einer bestimmten Bevölkerungsgruppe)²⁹ und dann erst wird analysiert, welchen Einfluss die Maßnahme auf diese Entwicklung hatte (eher induktiv). Beide Richtungen können auch verbunden werden.

Aus Sicht der quantitativen Sozialforschung gelten **experimentelle Untersuchungen mit Kontrollgruppen** als Königsweg der Wirkungsanalyse, u. a. da hiermit solide Kosten-Wirkungsanalysen möglich sind, die es erlauben, verschiedene Maßnahmen in ihrer Effizienz und Effektivität gut zu vergleichen.³⁰ Die Bildung solcher vergleichbaren Kontrollgruppen (Indivi-

²⁸ Die GTZ (2004, S. 10) hat für den Schritt vom direkten Nutzen einer Intervention (bei der Zielgruppe) bis zum indirekten Nutzen den irreführenden Begriff der „Zuordnungslücke“ geprägt. Je weiter Wirkungen in einer Wirkungskette von einer Intervention entfernt sind, desto wichtiger werden empirisch bestätigte Wirkungshypothesen, um diese Wirkungen zumindest plausibilisieren zu können.

²⁹ Hierfür wird teilweise der etwas irreführende Begriff „Bruttowirkung“ verwendet.

³⁰ Zu Vorzügen von Kontrollgruppen vgl. Rossi et al. (2004, S. 233-264). Die European Evaluation Society (2007) trat dem Trend zu einer einseitigen Festlegung auf randomisierte,

duen ohne Maßnahmen) ist hierbei die größte akademische und praktische Herausforderung; außerdem verlangt das Vorgehen eine genaue Planung schon zu Beginn der Entwicklungsmaßnahme. Das Vorgehen kann zudem anderen entwicklungspolitischen Grundsätzen entgegenstehen, z. B. dem Grundsatz der aktiven Partizipation der Zielgruppen bei der Durchführung der Entwicklungsmaßnahme. Oft erschweren auch die Komplexität der Maßnahmen und die Vielfalt der anvisierten Wirkungen einen solchen Analyseweg. In der Praxis der Entwicklungszusammenarbeit ist die Anwendung von Kontrollgruppen daher aus praktischen und ethischen Gründen in vielen Fällen nicht möglich oder sehr aufwändig.

In die gleiche methodische Gruppe – und damit mit ähnlichen praktischen Schwierigkeiten verbunden – fällt der **Vergleich ähnlicher Fälle** (z. B. werden vergleichbare Regionen, in denen eine bestimmte Entwicklungsmaßnahme nicht stattfindet, mit Regionen verglichen, in denen die Maßnahme ergriffen wurde).

Als Ausweg wird daher oftmals der **„Vorher-nachher“-Vergleich** herangezogen, bei dem die Situation nach Durchführung der Maßnahme mit der davor verglichen wird. Die meisten gebräuchlichen Verfahren unterliegen im Prinzip diesem Konzept. Die Hauptschwierigkeit besteht darin, dass exogene Faktoren hier nicht wissenschaftlich einwandfrei von den ‚Nettoeffekten‘, d. h. den Effekten, die tatsächlich auf die Maßnahme zurückzuführen sind, getrennt werden können. Hier ist nur eine mehr oder weniger fundierte Annäherung möglich. Zudem sind auch hier *baseline*-Untersuchungen zu Beginn der Maßnahme notwendig, andernfalls müssen *baseline*-Daten mühsam – und dadurch oft auch fehlerbehaftet oder unvollständig – rekonstruiert oder retrospektiv erhoben werden. Zudem werden bei einem solchen Vergleich ‚konservierende Wirkungen‘ nicht sicher entdeckt. Häufig wird untersucht, inwiefern die Ziele einer Maßnahme oder allgemeine Standards (z. B. im Bereich Gesundheit) erreicht worden sind. Ohne einen Vergleich mit *baseline*-Daten und ohne die Beachtung anderer möglicher Ursachen handelt es sich dabei jedoch nicht um einen Wirkungsnachweis.

Über diese Vergleiche hinaus gibt es weitere Möglichkeiten, Wirkungen zu analysieren: Bei der **qualitativen Pfadanalyse** (*process tracing*) wird die kausale Wirkungskette untersucht und in immer kleinere Schritte unterteilt. So können alternative Ursachen für bestimmte Wirkungen im Detail untersucht und zum Teil ausgeschlossen werden.

kontrollierte Studien (*randomised controlled trials*) mit der Forderung nach methodischer Vielfalt entgegen.

Wenn Wirkungen nicht direkt untersucht werden können, können auch **empirisch bestätigte Wirkungstheorien**³¹ als Hilfsmittel verwendet werden, die mit Hilfe anderer Fälle entwickelt wurden.³² Ein ähnliches Vorgehen ist die Bewertung der inneren Logik (logisches Wirkungsmodell³³) und der Qualität einer Entwicklungsmaßnahme (ihrer Planung und Durchführung), unter der Annahme, dass „qualitativ hochwertige“ Maßnahmen gute Wirkungen erzielen.³⁴

Schließlich können mögliche Nutznießer und andere Beteiligte oder „Experten“ nach ihrer (**explizit subjektiven**) **Ansicht bezüglich Wirkungen** gefragt werden, was auch retrospektiv möglich ist. Hier werden die Erfahrungen oder Einschätzungen von Expert/-innen oder Nutzer/-innen über die Entwicklung, wie sie sich mit der Maßnahme im Vergleich zur Situation ohne Maßnahme entwickelt hat, herangezogen (Schattenkontrollen). Oder die allgemein als „normale“ Entwicklung angesehene Situation wird implizit als Vergleichsgröße herangezogen. Die unterschiedlichen Wirkungsaussagen verschiedener Befragter können daraufhin miteinander verglichen und im qualitativen Sinne analysiert und bewertet werden (generische Kontrollen).

Meist ist eine Kombination unterschiedlicher Methoden zum Wirkungsnachweis hilfreich. Bei allen Nachweismöglichkeiten ist es sinnvoll, weiterhin die kontrafaktische Frage zu stellen, „Was wäre gewesen, wenn die Entwicklungsmaßnahme nicht stattgefunden hätte?“, und diese sorgfältig argumentierend möglichst gut zu beantworten.

³¹ Vgl. hierzu z. B. Hancock, J. u. a. (2003), die eine Klassifizierung der Anwend- und Übertragbarkeit von Innovationen vornehmen, darunter auch *best practice* (Wirksamkeit in unterschiedlichen *settings* belegt) und *policy principle* (Wirksamkeit quantitativ / wissenschaftlich und wiederholt bewiesen).

³² Ein Beispiel hierfür sind die von *Reflecting on Peace Practice* induktiv erarbeiteten Kriterien zur Wirksamkeit von Friedensförderung auf der Makroebene, siehe Anderson und Olson (2003).

³³ Vgl. z. B. Beywl (2006), S. 33.

³⁴ Für einen empirischen Nachweis des Zusammenhangs zwischen Qualität von Projektplanung und -management und Projektwirkung siehe z. B. Schubert, B.; Agrawal, R. C.; Böttcher, D. u. a. (1984).

1.3 Praktische Herausforderungen

*Wirkungserfassung im Rahmen von Monitoring und Evaluierungen*³⁵

Die Wirkungserfassung einer Entwicklungsmaßnahme erstreckt sich über deren gesamten Zyklus und muss daher von Anfang an eingeplant werden.

Im Idealfall sollte ein Entwicklungsvorhaben über eine *baseline*-Studie und ein wirkungsorientiertes Monitoringsystem verfügen. Dieses sollte auf bereits in der Planungsphase formulierten Wirkungshypothesen basieren, die eine der Grundlagen für eine Wirkungsanalyse darstellen können, die aber auch im Rahmen derselben auf ihre Relevanz hin überprüft werden müssen. *Baselines* sollten eher breit angelegt werden, da nicht immer genau vorhersehbar ist, welche Informationen später benötigt werden. In einem instabilen Kontext (z. B. in Konfliktgebieten mit starken Migrationen) kann ihre Verwendbarkeit allerdings beschränkt sein.

Wirkungsorientierte Monitoringsysteme sollten eine Kombination aus quantitativen und qualitativen Indikatoren beinhalten, die sich gegenseitig ergänzen. „Mittler und Zielgruppen sollten an der Entwicklung von Indikatoren beteiligt werden [...]“.³⁶ Die Durchführungsorganisationen sollten das Personal des Projektträgers/Partners (Führungsebene und Mitarbeiter/-innen) bei der Entwicklung des wirkungsorientierten Monitoring von Anfang an einbinden, um dessen Akzeptanz zu erreichen.³⁷ Förderorganisationen sollten ihre Partnerorganisationen dabei unterstützen, die eigenen z. T. bewährten oder nur ansatzweise erprobten Methoden der Planung, des Monitoring und der Evaluierung verstärkt auf Wirkungen auszurichten und weiterzuentwickeln.

In der Praxis sind diese Voraussetzungen nicht immer erfüllt. Dann ist es erforderlich, die vorhandenen Defizite nachträglich zu beheben, soweit es möglich ist. Bei Fehlen einer *baseline*-Studie kann versucht werden mit anderen Mitteln die Ausgangssituation zu rekonstruieren. Dies ist u. a. mit retrospektiven Verfahren möglich. Auch Wirkungshypothesen können nachträglich formuliert bzw. konstruiert werden. Hier besteht allerdings die Gefahr, dass diese ausschließlich auf die geplanten Wirkungen fokussieren und durch diese Verengung des Blickwinkels sonstige Veränderungen zu wenig berücksichtigt werden.

³⁵ Der Schwerpunkt liegt hier auf der Wirkungsanalyse, daher werden viele Aspekte, die im Zusammenhang mit Evaluierungen allgemein relevant sind, hier nicht behandelt.

³⁶ GTZ (2004), S. 19

³⁷ vgl. dazu auch: BMZ (2003): Deutsche Entwicklungszusammenarbeit. Kurzfassung der Querschnittsanalyse „Monitoringinstrumente in der deutschen EZ“, S. 3.

Die Zielformulierung einer Entwicklungsmaßnahme, die eigentlich die Ebene(n) der Nutzung und des direkten Nutzens darstellen soll, enthält in der Realität nicht immer eine Beschreibung der angestrebten Wirkungen.³⁸ Längst nicht in allen Entwicklungsmaßnahmen und -organisationen wurden bisher aussagekräftige und handhabbare Wirkungsmonitoringsysteme eingeführt. Die Bandbreite der Probleme reicht von computergestützten Systemen, die zum Teil große Mengen nicht zuverlässiger quantitativer Daten verarbeiten, bis hin zum vollständigen Fehlen quantitativer Daten und *baseline*-Studien. Nicht überall ist das Personal des Partners auf den verschiedenen Ebenen bei der Entwicklung des Wirkungsmonitorings ausreichend eingebunden. Auch die Beteiligung von Mittlern und vor allem von Zielgruppen (Endnutzern bzw. -begünstigten / *beneficiaries*) an der Entwicklung von Indikatoren wird teilweise noch immer vernachlässigt.

Die Definition und Einbeziehung von Kontrollgruppen in Wirkungsmonitoringsysteme ist aufwändig und teuer und oftmals sehr schwierig. Solen Kontrollgruppen bei einer Wirkungsanalyse (z. B. im Rahmen einer Evaluierung) einbezogen werden, müssen solche zunächst gefunden oder konstruiert werden. Sie müssen unter den gleichen Rahmenbedingungen leben und von den gleichen Voraussetzungen ausgehen wie die Zielgruppen, was oft zusätzliche Untersuchungsschritte erfordert. Je größer und komplexer Programme sind im Hinblick auf alle Kriterien und die anvisierten Wirkungen, umso schwieriger ist dies zu realisieren und manchmal ist es unmöglich. In vielen Fällen stehen dem auch ethische Gründe entgegen, denn wem könnte man zumuten als Kontrollgruppe zu fungieren, obwohl er von keinerlei Maßnahmen zur Verbesserung der eigenen Situation profitiert.³⁹

Bezüglich Langzeitwirkungen kann es u. U. schwierig sein, ausreichend informierte Gesprächspartner zu finden, wenn nur ein kleiner oder nicht mehr aufzufindender Personenkreis über die notwendigen Informationen verfügt.

Wahl eines geeigneten Verfahrens der Wirkungsanalyse

Entscheidend für die Wahl des Untersuchungsverfahrens ist der **Zweck der geplanten Wirkungsanalyse**, da dieser die Fragestellung mitbestimmt. Die

³⁸ Ebd., S. 10.

³⁹ Ein besonders anschauliches Beispiel hierfür wäre die Wirkungsanalyse von Maßnahmen zur Eindämmung von Gewalt gegen Gefangene, wie sie z. B. das IKRK durchführt. In der EZ kann es – auch wenn Haushaltsbefragungen oder auch Volkszählungen in fast allen Ländern gängig sind – der Bevölkerung mitunter schwer zu vermitteln sein, dass sie sich z. B. bezüglich Gesundheitsdiensten, von denen sie selbst nicht profitiert, dennoch an Untersuchungen der Entwicklungsorganisation beteiligen soll,

externen Aufgaben von Wirkungsanalysen (wie allgemein von Monitoring und Evaluierung) bestehen in der Rechenschaftslegung gegenüber dem (Auftrag-) Geber, dem Beitrag zum Wissensmanagement der Förder- und Durchführungsorganisation, sowie u. U. der Öffentlichkeitsarbeit und dem Marketing. Die internen Funktionen beinhalten die Bereitstellung von Informationen zur Projektsteuerung und zum Lernen auf der Durchführungsebene.⁴⁰ Daher ist eine Einbeziehung der Partner in die Planung, Durchführung und Auswertung von Wirkungsanalysen notwendig, was zur Erhöhung der Akzeptanz beiträgt.⁴¹ Um die Lernfunktion zu stärken, ist auf Geberseite zudem ein Kooperationsstil erforderlich, der von Transparenz und Partizipation geprägt ist. Um auch negative Erfahrungen nutzen zu können, sind auf gesellschaftlicher und institutioneller Ebene Transparenz und Meinungsfreiheit (Lernkultur) Voraussetzungen.⁴² Wenn auf Grund der politischen Rahmenbedingungen eine freie Meinungsäußerung nicht möglich ist, schränkt dies die Wahl der Verfahren ein.⁴³

Für die Partnerorganisationen ist die Konfrontation mit den verschiedenen Verfahren der Wirkungsanalyse unterschiedlicher Geber oft eine Herausforderung. Daher sind, wenn möglich, gemeinsame Analysen (im Rahmen von *joint evaluations*), z. B. in einem Sektor vorzuziehen, was jedoch einen nicht unbeträchtlichen Aufwand hinsichtlich der Abstimmung der verschiedenen Beteiligten erfordert.

Handelt es sich um eine Selbstevaluierung, muss sich das Verfahren an den Zielen und Kenntnissen der jeweiligen Organisation im Bereich Wirkungsanalyse orientieren, bzw. es muss externe Unterstützung gesucht werden.

Die Vorgehensweise wird sich auch unterscheiden müssen, je nachdem um welche Maßnahmentypen es sich handelt (Projekt, Programm, Strukturförderung von Organisationen, Fachberatung, Dialog, Politikberatung, Budgetfinanzierung usw.) und welche Zielgruppen sich definieren lassen. Welches Untersuchungsverfahren geeignet ist, hängt weiterhin von den **zeitlichen, personellen und finanziellen Ressourcen** ab, die zur Verfügung stehen. Eine Abwägung von Kosten und Nutzen ist hier erforderlich. Qualitative Methoden, die subjektive Wahrnehmungen einbeziehen, sind oft ausagekräftiger und kosteneffektiver als die Erfassung rein quantitativer Da-

⁴⁰ Vgl. GTZ (2004), S. 21-22.

⁴¹ AGKED/Misereor (1992), S. 16 fordert: „Für ein Evaluierungskonzept der kirchlichen Entwicklungsarbeit müssen [...] beiderseits akzeptable Methoden und Vorgehensweisen entwickelt und angewandt werden, die dem Erfordernis einer gemeinsamen Wirkungsbeobachtung mit unseren Partnern gerecht werden.“

⁴² Arbeitskreis „Evaluation von Entwicklungspolitik“ in der DeGEval (2005), S. 5.

⁴³ Vgl. dazu Englert (2006).

ten. Bezüglich des Nutzens und der Wirkungen des Vorhabens können die Zielgruppen meist sehr wertvolle Informationen liefern und sollten daher immer dazu befragt werden. Dabei sollte zunächst generell nach Veränderungen gefragt werden und nicht sofort nach dem Beitrag der Entwicklungsmaßnahme.⁴⁴

Um im konkreten Fall ein geeignetes Verfahren der Wirkungsanalyse auswählen zu können, ist der Zugang zu Informationen über möglichst viele verschiedene Verfahren erforderlich. Die Vorgaben der Auftraggeber sowie die zur Verfügung stehende Zeit schränken die Art der anwendbaren Verfahren jedoch ein. Einige Auftraggeber verlangen z. B. Kosten-Nutzen-Analysen und quantifizierbare Daten, während andere mehr an qualitativen Daten interessiert sind. Ideal ist methodische Offenheit und die Möglichkeit, verschiedene Verfahren zu kombinieren im Sinne eines Baukastenprinzips (vgl. Kapitel 1.2).

Aus Kostengründen besteht teilweise eine Tendenz, möglichst viele Maßnahmen nur oberflächlich auf Wirkungen hin zu analysieren, was zu einem geringen Erkenntnisgewinn führt und oft auch den Vorhaben nicht gerecht wird. Bei begrenztem Budget ist daher weniger oft mehr.

Praktische Herausforderungen für die Evaluierer/-innen

Eine wesentliche Herausforderung für die Evaluierer/-innen besteht darin, im Rahmen einer Evaluierung vor Ort, die je nach Auftraggeber und Umfang der Entwicklungsmaßnahme 1-3 Wochen dauert und während der oft nur wenige Tage für die Erfassung und Analyse von Wirkungen zur Verfügung stehen, die richtige **Balance zwischen der notwendigen Tiefe** (im Sinne einer wissenschaftlichen Ansprüchen genügenden Hypothesenprüfung) **und Breite** (im Sinne der möglichst weitgehenden Erfassung von relevanten Veränderungen) zu finden. Es sollen z. T. komplexe Wirkungshypothesen überprüft sowie gleichzeitig unbeabsichtigte Wirkungen und äußere Einflussfaktoren berücksichtigt werden. Letztere finden oft zu wenig Beachtung, da zur Überprüfung der Wirkungshypothesen ausreichend große Stichproben bei Zielgruppen und evtl. auch Kontrollgruppen erforderlich sind, was sehr zeitintensiv ist. Die Erfahrung zeigt, dass externe Wirkungsanalysen (etwa im Rahmen von Evaluierungen) üblicherweise bezüglich ihrer zeitlichen und finanziellen Ressourcen erheblichen Beschränkungen unterliegen, die sich zwangsläufig im analytischen Design niederschlagen. Insgesamt werden oft sehr hohe Anforderungen an die Evaluierer/-innen ge-

⁴⁴ GTZ (2004), S. 20; Neubert (2004).

stellt, die innerhalb des zur Verfügung stehenden Zeitrahmens nicht immer erfüllt werden können.⁴⁵

In der Regel besteht ein **Wirkungsanalyse-Team** aus zumindest je einem/einer Evaluierer/-in aus dem Geber- und dem Partnerland. Das stellt eine Chance dar, denn die lokalen Evaluierer/-innen können als kulturelle Brücke dienen, durch die manche Zusammenhänge, die sich durch den fremdkulturellen Kontext ergeben, schneller und besser deutlich werden, während entsandte Evaluierer/-innen nicht nur Erfahrungen aus anderen Zusammenhängen einbringen, sondern auch mit externem Blick „Selbstverständlichkeiten“ leichter in Frage stellen können. Eine gemeinsame Abstimmung des Verfahrens, das zur Anwendung kommen soll, setzt voraus, dass alle Mitglieder des Teams die notwendigen Informationen und eine geteilte Verantwortung haben.⁴⁶

Um die notwendigen Untersuchungen durchführen zu können, ist mitunter auch die Einbeziehung von einheimischem Personal oder von Mitarbeiter/-innen, die in der Durchführung der Entwicklungsmaßnahme tätig sind, als „Türöffner“ und notfalls auch als Interviewer/-innen notwendig.⁴⁷ Sie ist auch insofern sinnvoll, als sie die Aneignung und Nutzung der Ergebnisse durch die Partnerorganisation bzw. das Partnerland fördern kann. Insbesondere bei Partnern ohne Erfahrung in der Wirkungsanalyse ist dieses Personal jedoch oft nicht in Analyseverfahren ausgebildet. Dies verlangt vor der Datenerhebung häufig eine entsprechende Schulung und kommt damit *Capacity-Building*-Zielen entgegen. Bei der aktiven Einbeziehung von Akteuren, die in die Durchführung involviert sind, ist jedoch darauf zu achten, dass die Unabhängigkeit der Wirkungsanalyse nicht beeinträchtigt wird (*bias problem*). Auch die Qualität der Arbeit (speziell der Datenerhebung) ist bei soeben erst ausgebildetem Personal nicht immer sichergestellt.

Gegen externe Wirkungsanalysen, wie auch gegen den Aufbau von wirkungsorientierten Monitoringsystemen, bestehen in Partnerländern und -organisationen oft Vorbehalte, weil diese häufig noch nicht bekannt sind und Befürchtungen sowohl hinsichtlich der Ergebnisse als auch der Konsequenzen bestehen, die die Geberorganisationen ziehen könnten. Hier sollte sich eine neue Kultur durchsetzen, in der das Ziel darin besteht, negative Wirkungen möglichst früh und umfassend zu erkennen, um rechtzeitig ge-

⁴⁵ Vgl. dazu auch Englert (2006). Ein Problem besteht auch darin, dass einige Auftraggeber keine oder zu geringe Kenntnisse in empirischer Sozialforschung besitzen um realistische ToR erstellen zu können.

⁴⁶ Häufig trägt der oder die vom Geber beauftragte und entsandte Gutachter/-in die alleinige Verantwortung für das Gelingen der Wirkungsanalyse.

⁴⁷ Unabdingbar ist dies z. B. in Fällen, in denen für die lokalen Sprachen keine unabhängigen Übersetzer gefunden werden können.

gensteuern oder kompensieren zu können. Dies bedeutet, es sollte ein Einverständnis dafür geben, dass Fehler gemacht werden dürfen, soweit aus ihnen gelernt wird.

Neben den genannten Herausforderungen besteht bei einer nachträglichen Konstruktion von Wirkungshypothesen oder einer Rekonstruktion von *baseline*-Daten oft das Problem, dass auf Ergebnisse anderer Erhebungen und Evaluierungen zurückgegriffen werden muss, die Sachverhalte und Wirkungen in unterschiedlicher Gründlichkeit untersucht haben. Das Prinzip des sparsamen Umgangs mit Finanzmitteln erfordert zwar die Einbeziehung vorhandener Daten, jedoch ist deren kritische Analyse notwendig, um die Aussagekraft der gesamten Wirkungsanalyse nicht zu beeinträchtigen.

Die Evaluierer/-innen stehen bei der Planung des Untersuchungsdesigns vor der Entscheidung, in dem gegebenen engen Zeitrahmen entweder stärker den Erfordernissen einer hypothesengeleiteten Wirkungsanalyse genüge zu tun oder eine mehr explorative, induktive Vorgehensweise zu wählen, die stärker von den Wahrnehmungen der Zielgruppe wie auch anderer Beteiligter ausgeht und eher die Chance einer breiteren Wirkungserfassung eröffnet, d. h. auch ungeplante und extern verursachte Wirkungen zu erfassen: „Es genügt nicht, nur die Wirkungen des Vorhabens auf sein Umfeld zu beobachten. Auch die umgekehrte Blickrichtung vom Umfeld auf das Vorhaben ist von Bedeutung, weil sich Veränderungen in den Rahmenbedingungen oder Aktionen anderer Entwicklungsorganisationen positiv oder negativ auf die Zielerreichung auswirken können.“⁴⁸

Eine wichtige praktische Herausforderung stellt nicht zuletzt auch die **Vermittlung der Ergebnisse** von Wirkungsanalysen, insbesondere an nicht geschulte Personen, dar. Dazu gehört auch die klare, nutzer- und handlungsorientierte Formulierung von Empfehlungen. Hier bieten sich Verfahren zur Wirkungsanalyse an, die das gemeinsame Reflektieren der Ergebnisse durch die verschiedenen Akteure (Partnerorganisation, Zielgruppen, Geberorganisation) vorsehen.

Institutionelle Verankerung der Wirkungsanalyse

Grundsätzlich können Wirkungsanalysen als singuläres Ereignis (Einzelstudie) und unabhängig von der Ausrichtung einer Organisation durchgeführt werden. Wenn Wirkungsorientierung jedoch zum Management gehören soll, ist die Wirkungsanalyse institutionell zu verankern. Die Organisation von wirkungsorientiertem Monitoring und wirkungsorientierter Evaluierung sollte dann in einem einheitlichen und abgestimmten System erfolgen –

⁴⁸ GTZ (2004), S. 12.

Wirkungsorientierung von der Planung bis zur Evaluierung.⁴⁹ Schritte und Instrumente sollten aufeinander aufbauen, um Organisation und Mittelaufwand effizient und effektiv zu gestalten.

Die organisatorische Einführung der Wirkungsorientierung bezieht sich auf alle beteiligten Bereiche einer Organisation, die sich mit der Planung, Durchführung, Steuerung und Evaluierung von Entwicklungsmaßnahmen sowie organisationsinternem Lernen und Wissensmanagement beschäftigen. Grundlagen sind ein klares und verbindliches institutionelles Konzept sowie ein praxisorientierter und erprobter Leitfaden zur Umsetzung, in denen Begrifflichkeiten, Formate, Instrumente und Methoden für den Anwender erläutert werden.⁵⁰

Die Standardformate für Planung, Monitoring und Berichterstattung der jeweiligen Organisation müssen dem Konzept entsprechend formuliert werden. Partner und entsandtes Personal müssen geschult werden. Ein flexibles System zur Fortbildung vorhandenen Personals sowie zur Schulung neuen Personals oder neuer Partner muss eingerichtet werden.

Die Einführung der Wirkungsorientierung in Organisationen ist ein mehrjähriger Prozess, der erhebliche Personalkapazitäten benötigt. Dazu gehört auch die Förderung von Partnerorganisationen im Hinblick auf eine Entwicklung von Monitoringfähigkeiten, was gleichzeitig einen Beitrag zur Nachhaltigkeit lokaler Strukturen darstellt.

Die Evaluierungsabteilung der jeweiligen Organisation muss ihre *Terms of Reference* wirkungsorientiert und dem Konzept folgend erstellen. Zur partizipativen Wirkungserfassung vor Ort ist ausreichend Zeit einzuplanen. Methodischen Fragen muss in der Vorbereitung von Wirkungsanalysen angemessene Aufmerksamkeit gegeben werden. Zudem müssen ggf. externe Gutachter/-innen von der jeweiligen Organisation in den dort verwendeten Begrifflichkeiten und Methoden fortgebildet werden. Um diese Aufgaben kompetent erfüllen zu können, sollte die Evaluierungseinheit auch über Personal verfügen, das praktische Erfahrung in empirischer Sozialforschung besitzt.

Mittelaufwand

Der Aufwand an Finanzmitteln kann hier nur allgemein behandelt werden, es entstehen im Wesentlichen Personal-, Reise- und laufende Kosten; Materialkosten sind eher gering. Der Mittelaufwand richtet sich nach den institutionellen Rahmenbedingungen und dem Entlohnungsniveau. In allen Orga-

⁴⁹ Siehe auch Welthungerhilfe (2008a).

⁵⁰ Siehe Welthungerhilfe (2008b).

nisationen besteht somit die Notwendigkeit, bezüglich der Wirkungsanalyse ein **vertretbares Kosten-Nutzen-Verhältnis** zu erreichen. Es können jedoch keine festen Prozentsätze des Budgets für die Entwicklungsmaßnahme als Richtwerte angegeben werden, da die Maßnahmen zu unterschiedlich sind. Dies gilt allgemein auch für Evaluierungen. Die Kosten müssen jedoch immer in einem vertretbaren Verhältnis zum Volumen der Entwicklungsmaßnahme stehen.

Aufgrund der üblicherweise sehr eingeschränkten Verfügbarkeit von Zeit und Finanzmitteln besteht in der Praxis eine große Nachfrage nach Verfahren, die geeignet sind, Wirkungen mit vergleichsweise geringem Aufwand zu erfassen. Der Zeit- und Kostenaufwand für Wirkungsanalysen ist jedoch in jedem Fall nicht unerheblich und wird in der Praxis oft unterschätzt.

2. Wirkungsanalysen in der Praxis

Die verschiedenen Wirkungsanalyseverfahren in Kapitel 2.2 sind alphabetisch entsprechend ihrer offiziellen – oder am weitesten verbreiteten – Bezeichnung sortiert und folgen in ihrer Beschreibung einer durchgängigen Struktur. In Kapitel 2.3 folgen kurze Darstellungen von Ansätzen der Wirkungsanalyse, die nicht dem Muster der Verfahrenssteckbriefe folgen.

2.1 Struktur der Verfahrens-Steckbriefe

1. Allgemeine Informationen

Dieser einleitende Abschnitt enthält Informationen zur Genese des Verfahrens sowie Beispiele für seine praktischen Anwendungen in der Vergangenheit.

2. Anwendungsbereiche

Hier wird skizziert, für welche entwicklungspolitischen Sektoren, Interventionstypen und -ebenen das Verfahren besonders geeignet ist, zu welchen Zeitpunkten einer Entwicklungsmaßnahme und mit welchem Zeithorizont – Rückblick und/oder Prognose – es vorrangig eingesetzt werden kann, und ob Entwicklungsprozesse oder -zustände im Vordergrund der Betrachtung stehen.

3. Analyseansatz

In diesem Abschnitt wird Antwort auf die Frage gegeben, welche Wirkungen mit Hilfe des Verfahrens wie erfasst werden, das heißt welche Definition von Wirkung – oder welches Wirkungsverständnis – zugrunde liegt, ob bei der Analyse im Wesentlichen von der Ursache (der Entwicklungsmaßnahme) oder dem Kontext (den wirtschaftlichen, ökologischen oder soziokulturellen Veränderungen) ausgegangen wird, wie mit dem Zuordnungsproblem und konkurrierenden Ursachen für beobachtete Veränderungen oder Sachverhalte umgegangen wird und inwieweit der Wirkungsnachweis über subjektive Einschätzungen und/oder faktisch erfolgt.

4. Durchführung

Dieser Abschnitt ist wegen seines Umfangs in Teilabschnitte gegliedert: In a) Verfahrensablauf wird der (schrittweise) Ablauf der Wirkungsuntersuchung so knapp und gleichzeitig verständlich wie möglich beschrieben. In der Regel wird diese Erläuterung durch eine graphische Darstellung unterstützt. Die an der Wirkungsanalyse Beteiligten und ihre wesentlichen Rollen und Funktionen sind in b) Akteure dargestellt. In c) Datenerhebung und d) Datenauswertung werden die eingesetzten Methoden und Instrumente der Stichprobenauswahl, Informationssammlung und Datenverarbeitung und -validierung benannt und gegebenenfalls erläutert.

5. Nutzung der Ergebnisse

Hier wird beschrieben, in welcher Form (graphische Präsentation, Quantifizierungen, Ampeln und Noten usw.) und wie transparent und nachvollziehbar die Ergebnisse der Wirkungsanalyse dargestellt werden und inwieweit die Einspeisung der Ergebnisse in Prozesse der Steuerung, des Lernens und der Rechenschaftslegung im Verfahren angelegt ist.

6. Anforderungen

In diesem Abschnitt werden Hinweise auf den Durchführungsaufwand (Zeit und Kosten) sowie auf sonstige Voraussetzungen (Qualifikationen, Organisation) gegeben.

7. Besondere Eignung und Grenzen des Verfahrens

Hier wird eine abschließende Einschätzung des Wirkungsanalyseverfahrens vorgenommen.

Am Ende eines jeden Steckbriefs finden sich Hinweise auf eventuelle Urheberrechte, Quell- und weiterführende Literatur.

2.2 Verfahren der Wirkungsanalyse

Im Folgenden werden zunächst insgesamt 21 Wirkungsanalyseverfahren in Form von Steckbriefen vorgestellt:

- CEval-Ansatz zur Wirkungsevaluation / Stockmann'scher Ansatz
- Do-no-Harm (DNH)
- e-VAL®
- Impact Monitoring and Assessment (IMA) nach Herweg/Steiner
- Impact Pathway
- KAP- (Knowledge, Attitude and Practice) Studien
- Method for Impact Assessment of Programmes and Projects (MAPP)
- Most Significant Change (MSC)
- NGO-IDEAs Impact Toolbox
- Outcome Mapping (OM)
- Participatory Impact Monitoring (PIM) / FAKT
- Participatory Impact Monitoring (PIM) / SLE
- Participatory Livelihood System Analysis (PaLSA)
- Participatory Methodology for Measuring the Contribution of Volunteering to Development (MCVD)
- Peace and Conflict Assessment (PCA)
- Programm-integriertes Planungs-, Monitoring- und Evaluierungssystem (PriME) von InWEnt
- Quantitative experimentelle Designs
- Quantitative nicht-experimentelle Designs
- Real Time Evaluation (RTE) im Nothilfebereich
- Verbleibsstudien (Tracer Studies)
- Wirkungsorientiertes Monitoring (WoM) der GTZ

CEval-Ansatz zur Wirkungsevaluation / Stockmann'scher Ansatz

STEFAN SILVESTRINI / NICOLA READE

1. Allgemeine Informationen

Der im Rahmen zahlreicher Evaluationsstudien erprobte Ansatz zur Wirkungsevaluation von Projekten und Programmen wurde von Prof. Dr. Reinhard Stockmann am Centrum für Evaluation (CEval) entwickelt. Der Ansatz ist so gestaltet, dass er flexibel an verschiedenartige Themenfelder sowie an die jeweiligen organisatorischen oder institutionellen Rahmenbedingungen angepasst werden kann. Die theoretischen und methodischen Grundlagen werden zum ersten Mal 1996 in der Habilitationsschrift von Stockmann „Die Wirksamkeit der Entwicklungshilfe – Eine Evaluation der Nachhaltigkeit von Programmen und Projekten“ dargestellt. Des Weiteren enthält das 2006 ebenfalls von Stockmann veröffentlichte Buch „Evaluation und Qualitätsentwicklung – Eine Grundlage für wirkungsorientiertes Qualitätsmanagement“ eine detaillierte Ausarbeitung und Weiterentwicklung der Konzeption auf Grundlage der Erfahrungen aus der Evaluationspraxis sowie eine Handlungsanleitung zur Anwendung des Evaluationsansatzes.

Eine umfangreiche Sammlung zu praktischen Anwendungen in den Bereichen Entwicklungszusammenarbeit, Bildung, Umwelt, Arbeitsmarkt und Kultur findet sich auf der Website des CEval: www.ceval.de. Im „Handbuch zu Evaluation – Eine praktische Handlungsanleitung“ (Stockmann, 2007) werden darüber hinaus praktische Grundkenntnisse für die Planung, Durchführung und Anwendung des Evaluationsansatzes vermittelt.

2. Anwendungsbereiche

Aufgrund seiner Flexibilität eignet sich der Stockmann'sche Ansatz zur Wirkungsevaluation grundsätzlich für alle Interventionsmaßnahmen, die einen Projekt- bzw. Programmcharakter aufweisen. Es können somit einzelne Hilfs- und Kooperationsprojekte, Beratungs- und Trainingsmaßnahmen oder aber auch ganze Organisationen bzw. deren Programmlinien und -strategien

auf ihre Effizienz, Wirksamkeit und Nachhaltigkeit hin untersucht und bewertet werden.

Eine Einschränkung auf Sektoren ist ebenfalls nicht gegeben. Die meisten Anwendungserfahrungen liegen in der Entwicklungszusammenarbeit (technische, finanzielle und personelle Zusammenarbeit), in der Bildung (berufliche Bildung, Fort- und Weiterbildung, e-Learning und Hochschulwesen) und im Umweltbereich (Wasser- und Abfallwirtschaft, Handwerk, Umweltberatung) sowie in deren Querschnittsfeldern vor (ländliche Entwicklung, *capacity building* und industrielles Umweltmanagement in Entwicklungsländern etc.).

Bislang wurde der Stockmann'sche Evaluationsansatz primär zur formativen und summativen (sowohl Abschluss- als auch ex-post) Evaluation von Projekten und Programmen eingesetzt. So angewendet ermöglicht der Ansatz nicht nur eine Wirkungsbewertung, sondern darüber hinaus die Abschätzung der Nachhaltigkeit von Maßnahmenwirkungen sowie die Entwicklung von Handlungsempfehlungen zur weiteren bzw. zukünftigen Gestaltung der untersuchten bzw. weiterer geplanter Maßnahmen.

Der Stockmann'sche Evaluationsansatz kann aber auch ex-ante zur Überprüfung und Anpassung der Zielsystematik sowie der Maßnahmenkonzeption (Zielgruppenorientierung, Abschätzung nicht intendierter Wirkungen etc.) angewendet werden. Darüber hinaus eignet er sich zur Konzeption und Umsetzung von Monitoringsystemen, da mit der Anwendung des Evaluationsansatzes steuerungsrelevante Informationen und geeignete Indikatoren zur Erfolgsprüfung gewonnen und verarbeitet werden können. Durch die Erfassung von Ist- und Sollzuständen zu mehreren Erhebungszeitpunkten wird damit eine ganzheitliche Prozessbetrachtung und somit Effizienzbewertung ermöglicht.

3. Analyseansatz

Das Analyseraster ist auf die Erfassung sämtlicher relevanten Wirkungen (intendiert und nicht intendiert) auf der Mikro-, Meso- und Makroebene ausgerichtet und versucht damit alle durch die Interventionsmaßnahme hervorgerufenen (individuellen wie globalen) Struktur-, Prozess- und Verhaltensänderungen zu erfassen. Aufgrund der multidimensionalen Analysekriterien können dabei nicht nur die unmittelbar hervorgerufenen (direkten) Wirkungen sondern auch die nachgelagerten (indirekten) Effekte auf gesellschaftlicher bzw. systemischer Ebene identifiziert werden.

Die theoretische und methodische Konzeption zur wirkungsorientierten Evaluation von Projekten und Programmen beruht auf vier theoretischen Modellen: einem Lebenszyklus-Modell, einem organisationstheoretisch her-

geleiteten Wirkungsmodell, einem Innovations-Diffusions-Modell und einem Nachhaltigkeitsmodell. Während durch die Verwendung des Lebenszyklus-Modells der Prozesscharakter, die kausalen Verkettungen der einzelnen Phasen miteinander sowie mit der Wirkungsphase hervorgehoben werden, können mit Hilfe des kontingenztheoretischen Organisationsansatzes sowohl die inneren organisationalen Parameter als auch die äußeren situativen Bedingungen analysiert werden. Mit dem Innovations-Diffusions-Modell wiederum ist es möglich, die zentralen Faktoren zu identifizieren, die über die Verbreitung der durch die Projekte und Programme hervorgerufenen Veränderungen entscheiden. Die multidimensionale Nachhaltigkeitskonzeption erlaubt schließlich die Analyse der dauerhaften Ausbreitung von Programmwirkungen.

Kernelement der Konzeption ist ein aus dieser theoretischen Konzeption abgeleiteter Analyseleitfaden, in dem anhand von fünf Bewertungsbereichen alle relevanten Untersuchungsaspekte abgebildet werden:

1. Programm und Umwelt

Ausgangspunkt des Evaluationsverfahrens bilden Informationen über die Programm- bzw. Leistungsziele. Es werden die Programmkonzeption mit ihren Zielsetzungen, die implizite Programmtheorie (d. h. die angenommenen Ursache-Wirkungs-Zusammenhänge) sowie die Innovationskonzeption und die zu ihrer Umsetzung erforderlichen und zur Verfügung stehenden Ressourcen herausgearbeitet. Zur Identifikation des beabsichtigten Nutzens der Programminterventionen und Leistungen werden des Weiteren die Zielgruppen identifiziert und die gesellschaftlichen Bereiche, d. h. die Politik- und Praxisfelder, in denen die Programmwirkungen ihre Wirkungen entfalten sollen, beschrieben. Insbesondere im EZ-Kontext kommt dabei der Analyse der Ländercharakteristika eine bedeutende Rolle zu.

2. Planung und Durchführung

Gemäß des Lebenszyklus-Modells ist es zur ganzheitlichen Bewertung eines Programms erforderlich, den gesamten Planungs- und Implementierungsprozess in seinen einzelnen Phasen zu analysieren. Dementsprechend wird die Qualität des Programmverlaufs ebenfalls durch die Bewertung der Vorbereitung und Planung (Stakeholdereinbezug, Bedarfsorientierung, Machbarkeitsprüfung, Einschätzung der Leistungsfähigkeit der beteiligten Organisationen etc.), der Programmsteuerung während der Durchführung (Effektivität von M&E-Maßnahmen, Angemessenheit und Rechtzeitigkeit von Steuerungseingriffen und Anpassungsmaßnahmen), der Vorbereitung des Endes der Förderung (Vorhandensein und rechtzeitige Umsetzung adäquater Maßnahmen) sowie der Nachbetreuung bestimmt.

3. Interne Wirkungsfelder

Aufgrund des entscheidenden Einflusses der Leistungsfähigkeit der Organisation auf die Programmumsetzung und der damit erzielten Wirkungen werden die zentralen Parameter einer Organisation identifiziert: Die Ziele und ihre Akzeptanz, die Qualifikation des mit der Programmumsetzung betrauten Personals, die Organisationsstruktur (Arbeitsteilung, Koordination, Entscheidungsstrukturen, Kommunikation etc.), die finanziellen Ressourcen, die technische Infrastruktur sowie die technische Organisationskonzeption.

4. Externe Wirkungsfelder

Ein zentraler Bestandteil der Untersuchungsaspekte ist die Identifikation und Bewertung der Wirkungen bei den Adressaten, an die sich die Leistungen eines Programms richten. Dabei geht es nicht nur um die Erfassung des konkreten Nutzens für die Zielgruppen sondern auch um die Bewertung der Wirkungen bei den sozialen Gruppen, die nicht unmittelbar zu den Zielgruppen gehören sowie die nachgelagerten Wirkungen (Diffusion) sowohl in den gesellschaftlichen Subsystemen oder Politikfeldern, in denen die Interventionen stattfinden, als auch in denen, die mit ihnen in Zusammenhang stehen.

Bei der Untersuchung der Programmwirkungen werden insbesondere folgende Fragestellungen betrachtet:

- Welche Veränderungen sind aufgetreten? (Brutto-Wirkungen bzw. -Veränderungen)
- Welche dieser Wirkungen wurden durch das Programm verursacht? (Netto-Programmwirkungen)
- Welche der Programmwirkungen waren intendiert? (Programmziele)
- Inwiefern wurden die angestrebten Programmziele erreicht? (Soll-Ist-Vergleich)

5. Nachhaltigkeit

Schließlich ist die Nachhaltigkeit der erzeugten Wirkungen auf der Makro- und der Programmebene Gegenstand der Evaluation. Zum Einen wird also untersucht, inwieweit das Programm die wirtschaftlichen (Effizienz), sozialen (gesellschaftspolitische Relevanz) und ökologischen (Umweltverträglichkeit) Zielgrößen der Nachhaltigkeit erreicht. Zum Anderen werden die programmspezifischen Dimensionen der Nachhaltigkeit bewertet.

Ein weiterer Bestandteil des Stockmann'schen Evaluationsansatzes ist die Anwendung eines geeigneten Methodenmixes zur Datenerhebung. Da

im Rahmen von Evaluationsstudien die in der Regel erforderlichen experimentellen oder quasi-experimentellen Untersuchungsdesigns nicht umgesetzt werden können, wird dies durch die systematische Anwendung unterschiedlicher (qualitativer und quantitativer, empirischer und sekundäranalytischer, standardisierter und teil- bzw. nicht-standardisierter) Erhebungsmethoden kompensiert (vgl. 4c). Durch diese Kombination können die Untersuchungsergebnisse trianguliert werden, wodurch sie den wissenschaftlichen Kriterien der Validität, Reliabilität und intersubjektiven Nachvollziehbarkeit gerecht werden. Charakteristisch für den Evaluationsansatz ist schließlich dessen partizipative Ausrichtung, die einen intensiven Stakeholdereinbezug zu verschiedenen Zeitpunkten im Untersuchungsverlauf erfordert (vgl. 4b).

4. Durchführung

a) Verfahrensablauf

Entsprechend dem partizipativen Verständnis des Stockmann'schen Evaluationsansatzes beginnt die Durchführung einer Evaluationsstudie mit einem Abstimmungsworkshop mit dem Auftraggeber. Im Rahmen dieses Workshops werden die Untersuchungsziele präzisiert und die Konzeption der Studie sowie das methodische Vorgehen diskutiert. Im Anschluss an das Treffen sollten alle untersuchungsrelevanten Dokumente an das Gutachterteam übermittelt werden.

Der erste methodische Schritt besteht i. d. R. aus einer umfassenden Dokumentenanalyse. Dabei werden alle verfügbaren Informationsmaterialien sowie Dokumente zur Konzeption und zu Erfahrungen mit der Umsetzung des zu evaluierenden Programms gesammelt und ausgewertet. Hierbei wird oftmals das *Peer-Reviewing*-Verfahren angewendet, das zum Einen zur intersubjektiven Nachvollziehbarkeit der Analyseergebnisse beiträgt. Zum Anderen wird durch das Vier-Augen-Prinzip die vollständige Erfassung des Untersuchungsgegenstandes sichergestellt.

Ausgehend von den Untersuchungsfragen und auf Basis der Erkenntnisse des Abstimmungswshops sowie der Dokumentenanalyse wird ein Analyseraster erarbeitet, welches alle relevanten Untersuchungsaspekte umfasst. Dieses Raster wird im weiteren Verlauf der Studie mit dem Auftraggeber abgestimmt und kontinuierlich weiterentwickelt. Damit wird gewährleistet, dass sich die Evaluationsstudie an den aus Sicht des Auftraggebers bedeutsamen Fragestellungen orientiert und weitere empirisch gewonnene relevante Informationen in die Evaluationsergebnisse mit einfließen können. Anhand des Analyserasters werde im Folgenden die Erhebungsinstrumente (Interviewleitfäden, Fragebögen, Beobachtungsraster etc.) entwickelt.

Nach Abschluss der Instrumentenentwicklung beginnt die Datenerhebung, bei der die vorher entwickelten und getesteten Erhebungsinstrumente zum Einsatz kommen (vgl. 4c). Wie erwähnt wird hierbei insbesondere auf eine ausgewogene Kombination qualitativer und quantitativer Erhebungsinstrumente geachtet, um sowohl den Kriterien der Repräsentativität als auch der Genauigkeit der Untersuchungsergebnisse gerecht zu werden.

Nach Abschluss der Erhebungsphase werden alle qualitativen und quantitativen Daten einer synergetischen und interpretativen Auswertung unterzogen (vgl. 4d). Ziel der Analyse ist es, anhand empirisch fundierter Informationen Antworten auf die Fragestellungen zu liefern, weitere relevante Untersuchungsaspekte aufzuzeigen und zu diskutieren sowie praxisrelevante und umsetzbare Empfehlungen zu liefern. Die dabei gewonnenen Erkenntnisse münden in einen Berichtsentwurf, der dem Auftraggeber vorgelegt wird.

In ausreichendem zeitlichem Abstand zur Übermittlung des Berichtsentwurfs werden die Ergebnisse der Evaluationsstudie im Rahmen eines Abschlussworkshops vorgestellt und diskutiert. Dabei sollen Anmerkungen und ergänzende Hinweise angesprochen werden, um diese in die Endversion des Evaluationsberichts einfließen zu lassen. Die Endversion des Berichts enthält dabei neben den zentralen Untersuchungsergebnissen auch eine Beschreibung des Evaluationsdesigns und der eingesetzten Methoden, eine Dokumentation der eingesetzten Instrumente sowie ein zusammenfassendes *executive summary*, in dem die wichtigsten Ergebnisse und Empfehlungen zusammengefasst werden.

b) Akteure

Ein zentraler Aspekt des Evaluationsansatzes ist der intensive Einbezug der Stakeholder (Auftraggeber, Programmverantwortliche, Programmzielgruppen, Förderinstitutionen etc.) in den Evaluationsprozess. Wie in Abbildung 1 ersichtlich wird, werden die Evaluierten bereits in der Planungsphase der Evaluation im Rahmen der Operationalisierung des Untersuchungsgegenstandes eingebunden. Dabei sind sie an der Entwicklung von Untersuchungshypothesen beteiligt und liefern Vorschläge zur Optimierung des Untersuchungsdesigns. Während der Datenerhebung nehmen die Stakeholder in erster Linie ihre Rolle als Informationsgeber ein und tragen durch ihr Insiderwissen zur Identifikation weiterer relevanter Informationsquellen bei.

In der Phase der Datenanalyse wiederum werden sie im Rahmen der Rückkopplung der Untersuchungsergebnisse bei der Ergebnisbewertung und der Entwicklung von handlungsleitenden Empfehlungen beteiligt. Durch die Nutzung der fachspezifischen Expertise der Stakeholder (hier insb. der Programmverantwortlichen) wird sichergestellt, dass die im Evaluationsbericht

enthaltenen Bewertungen und Empfehlungen für alle Beteiligten nachvollziehbar sind und in der Praxis umsetzbar erscheinen.

ABBILDUNG 1: Darstellung der partizipativen Einbindung der Evaluierten in den Evaluationsprozess gemäß des Stockmann'schen Evaluationsansatzes



Quelle: Stockmann (1996), grafisch überarbeitete Darstellung

c) Datenerhebung

Wie unter 4a) geschildert, ist ein zentrales Charakteristikum des Stockmann'schen Evaluationsansatzes die Kombination unterschiedlicher Erhebungs- und Analyseverfahren. Dabei sind grundsätzlich alle quantitativen und qualitativen Erhebungsmethoden der empirischen Sozialforschung anwendbar, sofern sie dem Untersuchungsgegenstand gerecht werden und sie unter den Maßgaben der Umsetzungseffizienz und Zielgruppenangepasstheit sinnvoll erscheinen. Hierzu gehören u. a.:

- Standardisierte oder teilstandardisierte schriftliche Befragungen (postalisch oder online);
- leitfadengestützte Intensivinterviews und Expertengespräche;
- Gruppendiskussions- und -konsensverfahren (z. B. Delphi);
- teilnehmende und nicht-teilnehmende Beobachtungen.

Darüber hinaus werden i. d. R. auch interpretative bzw. hermeneutische Verfahren, wie beispielsweise Dokumentanalysen oder Auswertungen statistischer Daten, eingesetzt. Hierbei sind auch insbesondere Daten aus Monitoringsystemen und zurückliegenden Evaluationen von Interesse, die mittels des Analyseleitfadens einer sekundäranalytischen Untersuchung unterzogen werden können.

Als primäre Datenquellen kommen sämtliche Stakeholder eines Projekts oder Programms in Betracht, die entweder von dessen Wirkungen betroffen (nicht nur die Zielgruppen!) oder für dessen Umsetzung verantwortlich sind. Um hierbei möglichst objektive Ergebnisse erzielen zu können, muss auf eine Ausgewogenheit bei der Auswahl der Befragten hinsichtlich ihrer Interessenlagen bzw. individuellen Zielsetzungen geachtet werden. Die Größe des Samples definiert sich dabei nach der Anzahl der Stakeholdergruppen und ihrer Größe sowie nach dem für die Evaluation zur Verfügung stehenden Budget, nach Maßgabe der Angemessenheit für den Untersuchungsgegenstand. Zur Überprüfung bzw. Absicherung der Wirkungsordnung kommt i. d. R. bei der Datenerhebung, die ausschließlich von den Evaluatoren vorgenommen wird, ein Vergleichsgruppen- und, wo es möglich und ethisch vertretbar ist, ein Kontrollgruppendesign zur Anwendung.

Durch die partizipative Ausrichtung des Evaluationsansatzes (vgl. 4b) wird der Problematik der Wahrnehmungsunterschiede seitens der Stakeholder Rechnung getragen. Da mit der Zahl der Beteiligten auch die Zahl unterschiedlicher Sichtweisen zunimmt, die weder aus den untersuchungsleitenden Theorien ableitbar, noch aus dem Blickwinkel der Auftraggeber einer Evaluation offensichtlich sind, ist die Evaluation als interaktiver Prozess angelegt, der einen intensiven Dialog zwischen dem Evaluationsteam und den an der Evaluation und den zu evaluierenden Maßnahmen beteiligten Personen bzw. Institutionen erfordert. Durch die enge Abstimmung mit diesen Akteuren werden die verschiedenen Interessenlagen, Werte und Bedürfnisse berücksichtigt und darüber hinaus die Akzeptanz für die Durchführung und die Ergebnisse der Evaluation gesichert.

d) Datenauswertung

Die Datenauswertung wird wie die Datenerhebung ausschließlich von den Evaluatoren vorgenommen, wobei grundsätzlich sämtliche qualitativen und quantitativen Analyseverfahren der empirischen Sozialforschung, der Psychologie, Sozioökometrie oder der Wirtschaftswissenschaften (u. a.) eingesetzt werden können. Ausgangspunkt ist dabei immer das zugrunde liegende Datenmaterial, das im Hinblick auf seinen Informationsgehalt hinsichtlich der einzelnen Evaluationsfragestellungen, seine Standardisierbar-

keit (d. h. Anteil qualitativer und quantifizierbarer Informationen) und Analysierbarkeit im Vorfeld geprüft werden muss.

Insbesondere bei qualitativen Daten ist es angebracht, die Zuordnung von einzelnen Informationen zu den jeweiligen Untersuchungsaspekten im *Peer-Reviewing*-Verfahren vorzunehmen, um eine intersubjektive Nachvollziehbarkeit der Untersuchungsergebnisse sicherzustellen. Hierbei sollten die Gutachter unabhängig voneinander die Inhalte der Interviewtranskripte, Gesprächs- und Diskussionsprotokolle, Beobachtungs- und Fragebögen etc. den jeweiligen Fragestellungen zuordnen und eine Bewertung gemäß des zuvor festgelegten Analyserasters vornehmen.

Zur Konstruktion dieses Analyserasters, welches der Strukturierung der Datenerfassung und -auswertung dient (vgl. 4a), bietet der Stockmann'sche Evaluationsansatz einen umfassenden Evaluationsleitfaden, in dem umfassend potenziell in Frage kommende Untersuchungsaspekte aufgelistet und erläutert werden. Dieser Leitfaden kann flexibel dem jeweiligen Untersuchungsinteresse angepasst und im Verlauf der Untersuchung zur Überprüfung der Vollständigkeit der selbst entwickelten Fragestellungen eingesetzt werden. Darüber hinaus ist er als Hilfsmittel zur ‚Meta-Evaluation‘ von Evaluationskonzepten geeignet.

Bei der Interpretation der Evaluationsergebnisse werden wiederum die verschiedenen Stakeholdergruppen mit einbezogen. Ihre Aufgabe ist es, ihre Wahrnehmung der aus der Untersuchung gewonnenen Erkenntnisse zu diskutieren (z. B. im Rahmen eines Workshops) und eine eigene Bewertung vorzunehmen. Die aus den verschiedenen Perspektiven hervorgehende Gesamtbewertung wird wiederum von den Evaluatoren in den Evaluationsbericht mit aufgenommen und kritisch reflektiert.

5. Nutzung der Ergebnisse

Aufgrund der sehr hohen Transparenz und der umfassenden Partizipationsmöglichkeiten während des gesamten Evaluationsprozesses bestehen seitens der Evaluierten i. d. R. keine Probleme hinsichtlich der Nachvollziehbarkeit der Ergebnisse und Bewertungen. Da die gezogenen Schlussfolgerungen gemeinsam mit allen Beteiligten der Evaluation diskutiert werden, ist darüber hinaus von einer hohen Akzeptanz und damit Umsetzungswahrscheinlichkeit der getroffenen Empfehlungen zu rechnen. Im Rahmen der Diskussion wird den Beteiligten schließlich die Möglichkeit der Partizipation in Form einer eigenen Bewertung der Evaluationsergebnisse gegeben, die ebenfalls in das abschließende *Reporting* (Evaluationsbericht, Gutachten etc.) mit einfließt.

Erfahrungsgemäß nutzen die Auftraggeber die so umgesetzten Evaluationen zur Initiierung und Verstetigung von (organisationalen) Lernprozessen, zur Rechenschaftslegung gegenüber ihren Sponsoren sowie zur allgemeinen Verfahrensoptimierung im Rahmen laufender Programme und für die Konzeption zukünftiger Programme. Darüber hinaus werden Evaluationsergebnisse gelegentlich auch zur Öffentlichkeitsarbeit und zur strategischen Ausrichtung von Organisationsleitlinien verwendet. Grundsätzlich gibt es bei den potenziellen Nutzern der Evaluationsergebnisse keine Einschränkungen hinsichtlich der Hierarchieebenen oder funktionalen Aufgabenfelder innerhalb der auftraggebenden Organisationen.

6. Anforderungen

Der CEval-Ansatz zur Wirkungsevaluation stellt in zweifacher Hinsicht hohe Anforderungen an die Evaluationsbeteiligten. Zum Einen müssen die Evaluierenden über fundierte Kenntnisse in quantitativen und qualitativen Methoden der empirischen Sozialforschung verfügen. Zum Anderen erfordert der Ansatz aufgrund seiner partizipativen Ausrichtung eine aktive Beteiligung der Evaluierten bei der Planung und Durchführung des Evaluationsvorhabens und schließlich bei Bewertung der Untersuchungsergebnisse (vgl. 4b).

Hierdurch ergibt sich bisweilen ein im Vergleich zu manchen anderen Evaluationsansätzen etwas höherer Zeitbedarf, der jedoch durch die damit erreichte sehr hohe Validität und Reliabilität der Evaluationsergebnisse sowie der ebenfalls überaus zufriedenstellenden Ergebnisakzeptanz seitens der Evaluierten gerechtfertigt ist.

Ein Teil dieses zusätzlichen Bedarfs ist den durch die aktive Mitarbeit des Auftraggebers bzw. der Stakeholder vergleichsweise zeitintensiven Abstimmungsaktivitäten zuzurechnen, die je nach Evaluationsumfang und -dauer zwischen 3 und 10 Arbeitstagen (z. B. für Workshops oder Feedbackgespräche) in Anspruch nehmen können. Entsprechende personelle und zeitliche Ressourcen sind auch seitens des Auftraggebers mit einzuplanen.

Je nach Methodeneinsatz müssen für die Entwicklung des Analyserasters sowie die Ableitung der Erhebungsinstrumente in der Konzeptionsphase zwischen 3 und 8, bei sehr umfangreichen Befragungen (mündlich, schriftlich, online) verschiedener Stakeholdergruppen auch mehr Arbeitstage einkalkuliert werden. Der Zeitbedarf für die Dokumentenanalyse ist ebenfalls abhängig vom Umfang der in die Untersuchung einzubeziehenden Programmmaterialien und variiert demgemäß zwischen einem Arbeitstag und (insb. bei reinen *desk studies*) mehreren Wochen.

Die Dauer der Datenerhebungsphase ist ohne genaue Kenntnis der Untersuchungsgegenstandes und des Erkenntnisinteresses kaum abzuschät-

zen, da hierfür eine ganze Reihe von Faktoren (Untersuchungsort und Reisezeiten, Anzahl der zu interviewenden Personen, Umfang der standardisierten Befragung etc.) mit berücksichtigt werden muss. Das Gleiche gilt für die Datenauswertung und Berichtserstellung. Erfahrungsgemäß werden (mit einer erheblichen Streuung) in etwa zwischen 5 und 20 Arbeitstagen für die Datenerhebung und noch einmal die gleiche Summe an Tagen für die Analyse und das *reporting* (inkl. Workshop) benötigt.

Aufgrund seiner erheblichen Qualifikationsanforderungen für Evaluatoren hinsichtlich Methodenkompetenz, Moderations- und Mediationserfahrungen und *reporting skills* ist der Stockmann'sche Evaluationsansatz nur bedingt für Selbstevaluationen geeignet. Von einer Evaluation durch einen ‚Laien‘ ist auf jeden Fall abzusehen, da die Umsetzung der einzelnen Untersuchungsschritte von der Auftragsklärung bis zur Ergebnisrückkopplung eine professionelle Handhabung der einzelnen Erhebungs- und Auswertungsinstrumente erfordert, die nicht ohne eine entsprechende theoretisch-methodische und praktische Ausbildung leistbar ist. Bei einer Einbindung in ein kontinuierliches Monitoringsystem sollte zumindest in der Konzeptionsphase (zur Entwicklung des Analyserasters, der Konstruktion der Indikatoren und der Erhebungs- und Analyseinstrumenten) und zu Beginn der Datenerhebung ein professioneller Gutachter die Untersuchung begleiten. Zudem ist es ratsam, in geeigneten zeitlichen Abständen das Monitoringsystem auf seine Angemessenheit und Funktionsfähigkeit von einem Experten überprüfen zu lassen.

Seitens der Evaluierten ist darauf zu achten, dass die involvierten Mitarbeiter zur Weiterleitung von programmrelevanten Daten befugt sind und über eine ausreichende fachliche Kompetenz verfügen. Darüber hinaus stellt der Evaluationsansatz auch gewisse Anforderungen an die technische und organisatorische Infrastruktur des Auftraggebers. Des Weiteren ist es für die Umsetzung der Empfehlungen erforderlich, dass die Evaluation auf allen Hierarchieebenen ausreichend verankert ist. Das heißt, dass seitens des Auftraggebers sowohl auf der Sachbearbeiter- als auch der Leitungsebene entsprechende Ansprechpartner benannt werden.

7. Besondere Eignung und Grenzen des Verfahrens

Aufgrund der oben geschilderten Charakteristika wird das breite Anwendungsspektrum des Evaluationsansatzes deutlich. Durch seine Multiperspektivität eignet er sich hervorragend zur Initiation und Intensivierung von Lern- und Entwicklungsprozessen innerhalb von Organisationen aber auch zur Rechenschaftslegung und Legitimation von Entscheidungsprozessen. Die zahlreichen erfolgreichen Anwendungsbeispiele in unterschiedlichsten

Arbeitsgebieten (Entwicklungszusammenarbeit, Umwelt, Bildung etc.) belegen den praktischen Nutzen und die Übertragbarkeit des Verfahrens.

Rezension von Gerd-Michael Hellstern (in: Zeitschrift für Evaluation, Jg. 7, Heft 1, 2008 – S. 149-151): „Der CEval Ansatz einer wirkungsorientierten Evaluation erweist sich als fruchtbar für die Praxis. Er trägt der dualen Aufgabenstellung einer qualitativ anspruchsvollen Evaluation Rechnung, indem er zum einen Standards und Verfahren der Sozialwissenschaft einfordert sowie ein Verfahrensmodell entwickelt, das sich am Nutzen für den Auftraggeber orientiert und die Bedürfnisse der Zielgruppe betrachtet. Für Praktiker, aber auch für Wissenschaftler und Lehrende vermittelt die kompakte und nutzerfreundliche Darstellung der Vorgehensweise bei der Planung, Durchführung und Auswertung von Evaluationen eine gelungene Hinführung eines reflektierten Einsatzes von Evaluationsverfahren und -methoden. Die instruktiven Übersichten, Zusammenfassungen und praktischen Handreichungen leisten einen wichtigen Beitrag für eine professionelle Entwicklung der Evaluation. Im deutschsprachigen Raum, zum Nachschlagen und zur Orientierung wird das Handbuch für die Praxis und Lehre unersetzlich sein.“

Hervorzuheben ist, dass der Ansatz über ein solides theoretisches Fundament verfügt, das im Rahmen von Grundlagenforschungsarbeiten stetig weiterentwickelt wird. Die Verknüpfung der verschiedenen Analyseperspektiven (Lebenszyklus-Modell, organisationstheoretische Konzeption, Innovations- und Diffusionstheoretischer Ansatz, multidimensionales Nachhaltigkeitskonzept) erlaubt eine ganzheitliche Bewertung der Wirksamkeit und Effizienz sowie, im Falle von ex-post Evaluationen, der Nachhaltigkeit von Projekten und Programmen, die beispielsweise über einen reinen Stakeholder-orientierten Ansatz so nicht möglich erscheint. Die methodische Konzeption ermöglicht schließlich durch die systematische Anwendung sowohl qualitativer als auch quantitativer Erhebungs- und Auswertungsverfahren die Identifikation und Analyse von Ursachen-Wirkungs-Zusammenhängen.

Die Grenzen des Stockmann'schen Evaluationsansatzes sind derzeit sicherlich noch in seiner Anwendung im Bereich der ex-ante Evaluation zu sehen, da er bislang keine prognostischen Elemente umfasst und auch hierfür erforderliche Instrumente (Szenario-Techniken etc.) nicht Bestandteil der Konzeption sind. Zurzeit wird am CEval an einer diesbezüglichen Erweiterung der theoretischen und methodischen Grundlagen des Ansatzes geforscht.

Literaturhinweise

- Stockmann, Reinhard (2008): Evaluation and Quality Development. Principles of Impact-Based Quality Management. Frankfurt am Main: Peter Lang.
- Stockmann, Reinhard (Hrsg.) (2007): Handbuch zur Evaluation. Eine praktische Handlungsanleitung. Band 6 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann.
- Stockmann, Reinhard (2006a): Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Band 5 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann.
- Chinesische Übersetzung: Peking: China Social Science Press (in Vorbereitung), spanische Übersetzung: San José: Universidad de Costa Rica (in Vorbereitung).
- Stockmann, Reinhard (2006b): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Band 1 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann, dritte überarbeitete und aktualisierte Auflage.
- Auswahl von veröffentlichten Publikationen zu den Evaluationen (weitere Evaluationsbeispiele unter www.ceval.de):*
- Heinrich, Tina und Meyer, Wolfgang (2005): Entwicklung eines Monitoring-Systems für die politische Weiterbildung. In: Zeitschrift für Evaluation 3 (2), S. 271-291.
- Jakoby, Klaus-Peter; Schneider, Vera; Meyer, Wolfgang und Stockmann, Reinhard (2005): Nachhaltige Umweltberatung. Münster: Waxmann.
- Ludwig, Martina und Koglin, Ebba (2003): eBuT-Projektevaluation. In: dsv-Informationen, Vierteljahresschrift der Deutschen Vereinigung für Sportwissenschaft, 18. Jg. Hamburg: Deutsche Vereinigung für Sportwissenschaft.
- Stockmann, Reinhard (1996): Die Wirksamkeit der Entwicklungshilfe. Eine Evaluation der Nachhaltigkeit von Programmen und Projekten der Berufsbildung. Opladen: Westdeutscher Verlag.
- Stockmann, Reinhard; Caspari, Alexandra und Kevenhörster, Paul (2000): Langfristige Wirkungen der staatlichen Entwicklungszusammenarbeit. Ergebnisse einer Querschnittsevaluierung zur Nachhaltigkeit. Entwicklung und Zusammenarbeit 10.
- Stockmann, Reinhard; Meyer, Wolfgang; Krapp, Stefanie und Köhne, Godehard (2000): Wirksamkeit deutscher Berufsbildungszusammenarbeit. Ein Vergleich staatlicher und nicht-staatlicher Programme in der Volksrepublik China. Wiesbaden: Westdeutscher Verlag.
- Stockmann, Reinhard (2001): Evaluation entwicklungspolitischer Wirkungen in der staatlichen Zusammenarbeit. In: Hammerich, Kurz und Franke, Bettina (Hrsg.): Nord-Süd/Süd-Nord-Beziehungen. St. Augustin: Academia.
- Stockmann, Reinhard (2001): Evaluation der Nachhaltigkeit von Umweltberatung. In: Stockmann, Reinhard u. a. (Hg.): Umweltberatung und Nachhaltigkeit. Berlin: Erich Schmidt Verlag.
- Stockmann, Reinhard; Meyer, Wolfgang; Kohlmann, Uwe; Gaus, Hansjörg und Urbahn, Julia (2001): Nachhaltige Umweltberatung. Eine Evaluation von Umweltberatungsprojekten. Opladen: Leske + Budrich.

- Stockmann, Reinhard (2005): Qualitätsmanagement und Evaluation bei eLearning Programmen. In: Igel, Christoph und Daus, Reinhard (Hrsg.): Handbuch eLearning. Schorndorf: Hofmann.
- Stockmann, Reinhard und Krapp, Stefanie (2005): Evaluation des DAAD Programmbereichs I Stipendien für Ausländer. Bonn: DAAD.

Do No Harm (DNH)

MARTIN QUACK / DIRK SPRENGER

1. Allgemeine Informationen

Das Verfahren Do No Harm (DNH) geht auf ein Projekt verschiedener US-amerikanischer und internationaler NRO (ehemals bekannt als das Local Capacities For Peace Project) zurück, das in den frühen 1990er Jahren begann. Entscheidenden Anteil an der Entwicklung des Instruments hatte Collaborative for Development Action (CDA), eine Consulting-Firma in Cambridge, Massachusetts, insbesondere deren Präsidentin Mary B. Anderson. DNH soll dazu dienen, negative Wirkungen von Entwicklungsmaßnahmen auf gewaltträchtige Konflikte zu minimieren. Mit seinem einprägsamen Titel hat es sich international schnell verbreitet. In der deutschen EZ wird es u. a. von den folgenden Organisationen angewendet: Welthungerhilfe, Weltfriedensdienst, AGEH, World Vision Deutschland, Evangelischer Entwicklungsdienst, Peace Brigades International, DED.

2. Anwendungsbereiche

DNH wurde für die Nothilfe entwickelt, es gibt aber keine prinzipiellen Einschränkungen auf bestimmte Sektoren. Alle Entwicklungsmaßnahmen in gewaltträchtigen Konflikten (*working in conflict*) sowie Maßnahmen zur Friedensförderung (*working on conflict*) können mit DNH untersucht werden. Das Verfahren ist auf allen gesellschaftlichen Ebenen anwendbar, z. B. auf der Ebene von Organisationen. Der Fokus liegt jeweils auf der Ebene, auf der die Entwicklungsmaßnahme ansetzt. Im zeitlichen Ablauf eines Projekts kann DNH zu Beginn (Planung), im Verlauf (Monitoring) und am Ende sowie ex-post (Evaluierung) verwendet werden. Da einige der Untersuchungsgegenstände sich erst im Projektverlauf konkretisieren ist die Anwendung von DNH nach der Planung besonders wichtig. Der Schwerpunkt der Analyse liegt eher auf der Untersuchung von Prozessen (Transfer von Ressourcen und impliziten ethischen Botschaften) als von Zuständen.

3. Analyseansatz

Jede Maßnahme in einem Konfliktgebiet wirkt sich auf den Konflikt aus, ganz gleich in welchem Sektor sie arbeitet, sie wird Teil des Konflikt-Kontextes. Oft sind diese Wirkungen ungeplant und negativ. DNH untersucht die Auswirkungen von Ressourcentransfers und impliziten ethischen Botschaften von der Maßnahme auf den Konflikt. Sowohl die Wirkungen auf die *dividing factors* und die *sources of tension* in einem Konflikt als auch auf die *connecting factors* und die *local capacities of peace* werden untersucht. Mit Hilfe der Entwicklung von Handlungsoptionen sollen negative Wirkungen bewusst minimiert und positive Wirkungen verstärkt werden.

Mit dem Zuordnungsproblem wird nicht explizit umgegangen. Maßgeblich ist die subjektive Einschätzung der Akteure, die DNH durchführen (z. B. in projektinternen, partizipativen Workshops). Da es um die Minimierung von potenziell negativen und die Maximierung von potenziell positiven Nebenwirkungen (die Hauptwirkungen betreffen nicht den Konflikt) geht, ist der Nachweis von tatsächlich eingetretenen positiven oder negativen Nebenwirkungen zwar nicht notwendiger Bestandteil des Verfahrens aber dennoch sehr sinnvoll.

4. Durchführung

a) Verfahrensablauf

Die Zahl der einzelnen Schritte von DNH kann sich unterscheiden, die Logik ist in allen Varianten die gleiche:

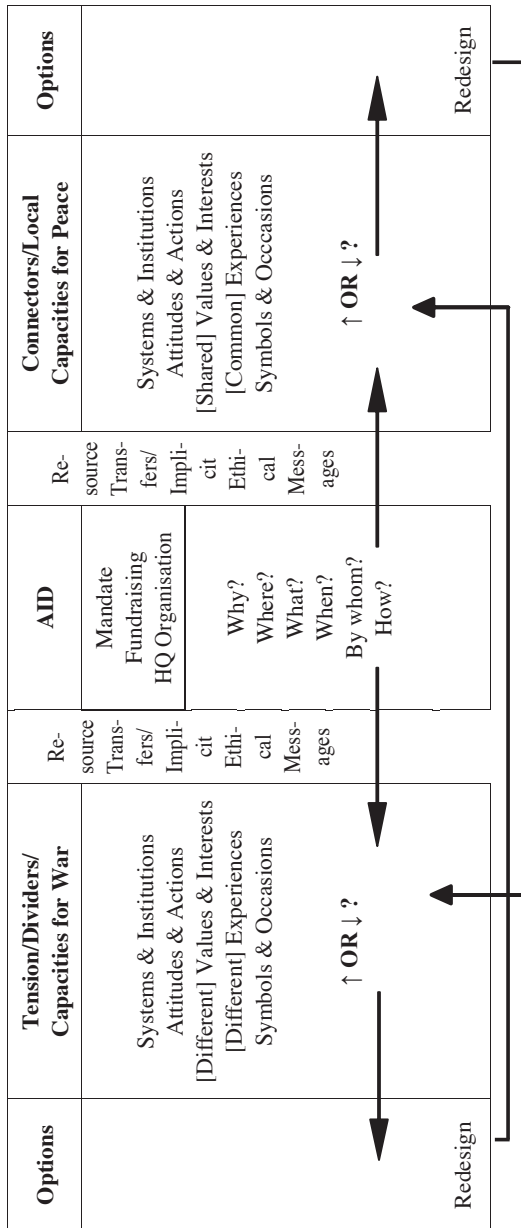
1. Zunächst wird in einer Kontextanalyse die angemessene „Arena“ identifiziert. Dabei handelt es sich um den geographischen und sozialen Raum, der für die zu untersuchende Maßnahme angemessen ist. In dieser Arena werden Konflikte zwischen Gruppen, die gewaltsam eskaliert sind oder eskalieren können, identifiziert.
2. Im zweiten Schritt werden die *dividing factors* und die *sources of tension* auf verschiedenen Ebenen analysiert: Systeme und Institutionen, Einstellungen und Verhalten, divergierende Werte und Interessen, divergierende Erfahrungen sowie Symbole und Gelegenheiten (*opportunities*).
3. Im dritten Schritt werden in analogen Dimensionen die *connecting factors* und die *local capacities for peace* analysiert: Systeme und Institutionen, Einstellungen, Verhalten, gemeinsame Werte und Interessen, gemeinsame Erfahrungen sowie Symbole und Gelegenheiten.

4. Die Untersuchung der Entwicklungsmaßnahme stellt den vierten Schritt dar. Dazu wird die durchführende Organisation bezüglich ihres Mandats, ihrer Finanzierung und bezüglich der Strukturen in ihrer Zentrale untersucht. Außerdem wird die Entwicklungsmaßnahme selbst mit Hilfe der folgenden Fragen analysiert: Weshalb (formuliertes Ziel), wo (Ort der Maßnahme), was (Leistung der Maßnahme), wann (Zeitraum), mit wem (Zielgruppen), durch wen (Mitarbeiter/-innen) und wie genau?
5. Im fünften Schritt werden die Wirkungen von Ressourcentransfer (RTs) und impliziten ethischen Botschaften (IEMs) untersucht:
 - RTs: Verteilungseffekte; Markteffekte; Substitutionseffekte; Legitimisierungseffekte;
 - IEMs: Kulturelle Effekte; Lebensstandard; Ressourcennutzung; Konkurrenz zwischen externen Akteuren; Hilflosigkeit; Angst, Spannung und Misstrauen; unterschiedlicher Wert von Menschen; Dämonisierung und *victimisation*; Waffen und Macht. Implizite ethische Botschaften gehen von problematischem Verhalten von Projektbeteiligten aus, das oft gar nicht projektbezogen ist (z. B. mangelnder Respekt für lokale Kulturen, Nutzung knapper Ressourcen für private Zwecke).

Es werden sowohl die (potenziellen) Wirkungen der RTs und IEMs auf *dividing factors* und *sources of tension* als auch auf *connecting factors* und *local capacities for peace* identifiziert. Dies geschieht entlang der Eigenschaften der Maßnahme (siehe vierter Schritt) und der identifizierten *dividers* und *connectors* (zweiter und dritter Schritt).

6. Wenn ein Element der Maßnahme negative Wirkungen auf *dividers* oder *connectors* hat werden im sechsten Schritt Alternativen entlang der Eigenschaften der Maßnahme (s. vierter Schritt) entwickelt.
7. Der abschließende siebte Schritt dient dem Test der Alternativen und dem *redesign* der Maßnahme. Dazu werden die potenziellen Wirkungen der Alternativen auf *dividers* und *connectors* abgeschätzt, die besten Alternativen werden ausgewählt.

ABBILDUNG 2: Übersicht über Do No Harm



Quelle: Nach Anderson (1999), S. 74

b) Akteure

Die Art der Durchführung ist flexibel, üblich sind partizipative Workshops mit Beteiligten der Maßnahme, in denen die einzelnen Schritte durchgeführt werden. Zielgruppen müssen nicht in jedem Fall beteiligt werden, oft ist dies jedoch geboten.

DNH ist in erster Linie für Selbstevaluierungen und das damit verbundene interne *capacity building* geeignet. Die Integration von DNH in externe Evaluierungen ist möglich und oft sehr sinnvoll (z. B. bei Sektor-evaluierungen in Krisengebieten).

c) Datenerhebung

Die wichtigsten Datenquellen sind die Workshop-Teilnehmer/-innen mit ihrem Wissen. *Sampling* ist nicht vorgesehen. Monitoringdaten können bei der Anwendung von DNH während oder nach Ende einer Maßnahme hilfreich sein. Als Instrumente der Datenerhebung dienen in erster Linie die Workshops mit Visualisierungen. Unterschiedliche Perspektiven/Wahrnehmungen werden in den Workshops implizit berücksichtigt.

d) Datenauswertung

Die Workshops mit verschiedenen Visualisierungen dienen auch als Instrumente zur Datenauswertung. Das Auswertungsraster ergibt sich aus der Verbindung der Kategorien der *dividers* und *connectors* mit den Eigenschaften der Maßnahme sowie mit den verschiedenen Ressourcentransfers und impliziten ethischen Botschaften (siehe 4a). Die Validierung der Ergebnisse geschieht implizit in den Workshops.

5. Nutzung der Ergebnisse

Die Ergebnisse von DNH sind für Workshop-Beteiligte transparent, für andere Akteure evtl. nur eingeschränkt. Eine Quantifizierung oder Aggregation der Ergebnisse ist nicht vorgesehen. Die Ergebnisse sind verständlich weil sie sehr konkret sind und grafisch unterstützt werden. Die Nutzung der Ergebnisse ist durch den letzten Schritt im Verfahren angelegt; sie dienen den Beteiligten vor allem zur Steuerung der Maßnahme.

6. Anforderungen

Der Zeitaufwand beträgt ca. 1 Tag, bei Updates im Rahmen des Monitoring ggf. weniger. Die Einführung des Konzeptes dauert ca. 1-2 Tage. Mit fortgeschrittenem Grad der Planung bzw. Implementierung ist eine externe Moderation angebracht. Kosten entstehen lediglich für die Moderation sowie für Material.

7. Besondere Eignung und Grenzen des Verfahrens

Die Stärke von DNH liegt in der Sensibilisierung dafür, dass Maßnahmen in Konfliktsituationen nie neutral sind. DNH nimmt auch Optionen für friedensfördernde Wirkungen von Maßnahmen in den Fokus. Es handelt sich um ein gut ausgearbeitetes Instrument, das von einer Vielzahl von Nothilfe- und EZ-Organisationen angewandt wird.

Eine Schwäche von DNH besteht in seiner Tendenz zur Bipolarität von friedensfördernden und konfliktverschärfenden Faktoren. Manche Wirkungen werden nur sichtbar, wenn darüber hinaus das Projekt und sein aktueller Kontext in den Fokus genommen werden. DNH lässt sich nicht direkt in andere Planungsinstrumente (z. B. *logframe*) integrieren, weil es mit anderen Kategorien arbeitet; dies ist eine weitere Schwäche.

Literaturhinweise

Anderson, Mary B. (1999): Do No Harm: How Aid can support Peace – or War. Boulder/ London: Lynne Rienner.

Anderson, Mary B. (Hrsg.) (2000): Options for Aid in Conflict: Lessons from Field Experience. Collaborative for Development Action. Cambridge, MA. (www.cdainc.com/publications/dnh/options/OptionsManual.pdf).

Zahlreiche weitere Dokumente zugänglich über www.cdainc.com.

1. Allgemeine Informationen

e-VAL[®] ist ein computergestütztes, qualitatives Interviewverfahren, das die Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) zusammen mit der Bremer Unternehmensberatung „nextpractice“ seit 2000 entwickelt hat. Es basiert auf ähnlichen Untersuchungsverfahren, die seit mehr als 10 Jahren auch in der Privatwirtschaft angewandt werden. Es dient der systematischen Erfassung und Analyse subjektiver Meinungen.

In der GTZ kommt e-VAL[®] bei Monitoring, Selbstevaluierungen und unabhängigen Evaluierungen zum Einsatz und stellt damit einen Input in eine umfassendere Betrachtung bzw. Evaluierung dar.

Da die Befragten im Interviewprozess selbst entscheiden, worüber sie Aussagen machen, kann e-VAL[®] nur dann für Wirkungsanalysen genutzt werden, wenn Wirkungen thematisiert werden. Befragte können beispielsweise in ihren Aussagen den Schwerpunkt auf Durchführungsthemen wie Zusammenarbeit mit Partnern, Strategie- oder Prozessgestaltung legen, ohne auf Wirkungen einzugehen. Wenn sie über Wirkungen sprechen, bilden ihre qualitativen Aussagen die Basis für weitergehende Wirkungsanalysen.

2. Anwendungsbereiche

e-VAL[®] wird von der GTZ in der Regel zum Ende einer Programm- oder Projektphase durchgeführt. In der Vorbereitung von unabhängigen Evaluierungen kommt es auch zu anderen Durchführungszeitpunkten der Entwicklungsmaßnahmen zum Einsatz. Die GTZ setzt es ohne sektorale Begrenzung für alle Typen von EZ-Projekten und -Programmen ein. Es wird dabei für Interventionen sowohl auf Makro- als auch auf Meso- und Mikroebene angewandt.

e-VAL[®] eignet sich zum Vergleich der Einschätzungen verschiedener Interessengruppen und von Zuständen (Situationen) zu bestimmten Zeitpunkten. Ebenso sind übergreifende Auswertungen auf Ebene von Sektoren, Multi-Geber-Programmen oder Länderportfolios möglich.

3. Analyseansatz

Das Verfahren beruht auf der Erfassung subjektiver Einschätzungen von Vertretern unterschiedlicher Interessensgruppen. Zu den Interessensgruppen zählen bei der Version 2.0 von e-VAL[®] die an der TZ-Maßnahme beteiligten Partner, die Mittler, die Zielgruppen, die GTZ-Mitarbeiter und andere Beteiligte. Mindestens 3 Interessensgruppen mit jeweils wenigstens 4 Interviewten sollen in einem e-VAL[®]-Prozess vertreten sein.

Die subjektiven Aussagen werden von den Interviewten in Unterschiedspaaren formuliert. Zuerst wird ein Bewertungsmaßstab und anschließend dazu ein Gegenpol formuliert (z. B.: „Dienstleister werden ohne klare Auswahlkriterien engagiert.“ vs. „Dienstleister werden durch einen neutralen Ausschuss unter Beachtung vordefinierter Kriterien ausgewählt“ oder „Im Projektgebiet herrschen konfliktive Beziehungen und Gewalt vor.“ vs. „Die sozialen Beziehungen haben sich im Projektgebiet verbessert“). Dieser Methodenschritt geht davon aus, dass Menschen Prozesse, Produkte oder Personen anhand von Unterschieden bewerten (Konstruktivismus).

Bei e-VAL[®] werden die zu evaluierenden Kriterien, Elemente genannt, durch die Interviewten anhand der von ihnen formulierten Unterschiedspaare bewertet. Die Elemente lassen sich grob in 4 Kategorien unterscheiden:

1. Entwicklung: Elemente, die Interpretationen zum zeitlichen Verlauf des Vorhabens ermöglichen können: z. B. Situation heute oder Situation nach Ende der Entwicklungsmaßnahme;
2. Erfolgsvorstellungen: Elemente, die einen Vergleich von Erfolgsvorstellungen unterschiedlicher Interessensgruppen ermöglichen, z. B. gewünschter Erfolg aus Sicht der Zielgruppen, gewünschter Erfolg aus Sicht der GTZ, Erfolg aus Sicht des Interviewten);
3. Beiträge: Elemente, die eine Analyse der Beiträge unterschiedlicher Gruppen zu der Entwicklungsmaßnahme ermöglichen (z. B. Beiträge des Partners X, Beiträge der Zielgruppen, Beiträge der GTZ);
4. Konzepte: 2 Elemente, die sich auf die Unterschiede von Konzepten des Vorhabens beziehen (Konzept am Anfang der Entwicklungsmaßnahme, Konzept heute).

Das Verbinden von subjektiven Aussagen mit vorhabensspezifischen Evaluierungskriterien (*repertory grid*) erlaubt es, in kürzester Zeit sowohl quantitative als auch qualitative Ergebnisse zu erhalten. Die Serverauswertung steht innerhalb weniger Minuten zur Verfügung.

Das Verfahren ermöglicht die Erfassung verschiedenster Arten von Wirkungen. Die interviewten Personen entscheiden über die Themen, die zur Sprache kommen, und damit darüber welche Wirkungen in die Betrachtung einbezogen werden. Grundsätzlich können sowohl positive wie negati-

ve, geplante wie auch ungeplante Wirkungen thematisiert werden. Im Interview selbst wird die Zuordnungsproblematik nicht thematisiert.

4. Durchführung

a) Verfahrensablauf

Vorbereitung

Im ersten Schritt trifft die auftragsverantwortliche Person die für die Anwendung von e-VAL[®] notwendigen Vorbereitungen im Projektbearbeitungssystem der GTZ. Im Gegenzug erhält sie ihre individuellen Zugangsdaten und den Link zum e-VAL[®]-Server. Danach wählt sie aus der Liste der hierfür von der GTZ qualifizierten Personen einen Evaluierungsbeauftragten aus, der den gesamten e-VAL[®]-Prozess steuert. Weitere Schritte sind dann:

- Auswahl der Interviewer und Auswerter;
- Inhaltliche Vorbereitung, u. a. Festlegung der Merkmale des Projekts/Programms, die bewertet werden sollen. Es handelt sich dabei um vorhabenspezifische Ausprägungen von Standardelementen. So variieren zum Beispiel je nach Entwicklungsmaßnahme die Partner, die Beiträge liefern, oder es kann anstelle des Elements „Situation am Beginn der Maßnahme“ ein anderer Vergleichszeitpunkt in der Vergangenheit gewählt werden, an dem zum Beispiel eine einschneidende Veränderung im Durchführungskonzept beschlossen wurde (z. B. nach einer Evaluierung);
- Auswahl der Interviewpartner und ihre Einordnung in die passenden Befragengruppen.

Interviewteil und Themenzuordnung

- Information der Interviewpartner über Funktion und Ablauf des e-VAL[®]-Interviews;
- Durchführung der Interviews durch speziell geschulte e-VAL[®]-Interviewer;
- Kategorisierung der gemachten Aussagen in Themenbereiche, um die Interpretation zu erleichtern und um sie auf evaluierungsrelevante Themen zu fokussieren.

Interpretation und Berichterstellung

- Interpretation der e-VAL[®]-Ergebnisse durch speziell ausgebildete Interpretatoren;
- Erstellung eines Interpretationsberichts.

Die e-VAL[®]-Ergebnisse werden für Steuerungsentscheidungen durch das betroffene Vorhaben, durch Evaluierer bei Vorbereitung und Durchführung einer auf e-VAL[®]-Ergebnissen aufbauenden Evaluierung sowie durch Organisationseinheiten der GTZ zur Durchführung von Querschnittsauswertungen genutzt.

b) Akteure

In der Regel werden folgende Gruppen befragt: Partnerpersonal, Zielgruppen-/Zivilgesellschaftsvertreter, GTZ-Mitarbeiter. Zusätzlich zu Zielgruppen können auch Mittler sowie andere Beteiligte (Gebervertreter o. ä.) befragt werden.

Die Auswahl der zu befragenden Personen erfolgt durch den Auftragsverantwortlichen der GTZ. Dabei sind jene Personen auszuwählen, die sich durch einen besonders guten Einblick in das zu evaluierende Vorhaben auszeichnen. Die Interviewpartner innerhalb einer Befragtengruppe sollten einen repräsentativen Querschnitt gegebener Interessen und Meinungen widerspiegeln, um die Aufnahme unterschiedlicher Perspektiven und Wahrnehmungen zu ermöglichen.

Die Interviews werden durch speziell geschulte Interviewer durchgeführt und durch dafür ausgebildete Interpretatoren ausgewertet.

c) Datenerhebung

Alle Personen werden direkt und einzeln interviewt. Dabei gibt es keine vorformulierten Fragen. Die Befragten sagen, was Ihnen wichtig ist. Das Interviewverfahren ist für die einzelne interviewte Person transparent. Die interviewte Person verfolgt und korrigiert, falls notwendig, die Dokumentation ihrer Aussagen, um eine wortgetreue Erfassung sicherzustellen. Erst wenn sie mit der erfassten Formulierung einverstanden ist, wird das Interview fortgesetzt.

Die Länge des Interviews wird durch die interviewten Personen bestimmt. Erst wenn sie sagen, dass sie alle für sie wichtigen Aussagen zu der Entwicklungsmaßnahme gemacht haben, wird dieser Interviewteil beendet.

Im Anschluss daran sieht der Interviewablauf vor, dass sich die interviewte Person mit dem Gesamtergebnis und den daraus berechneten Beziehungen der einzelnen Elemente des Interviews auseinandersetzt. Hierzu

wird ihr die grafische Auswertung ihres Interviews gezeigt, und ein Überblick über das Ergebnis der gemachten Bewertungen gegeben. Falls hierbei Widersprüchlichkeiten deutlich werden zwischen dem, was sie zum Ausdruck bringen wollte, und dem, was tatsächlich erfasst wurde, können Bewertungen korrigiert werden. Erst nach dieser Bestätigung des Ergebnisses durch die befragte Person wird das Interview abgeschlossen, gesichert und verriegelt. Hiernach gibt es keine Möglichkeit mehr, das Interview einzusehen, weiterzuführen oder zu korrigieren.

d) Datenauswertung

Die Datenauswertung erfolgt durch einen zentralen Server. Alle geführten Interviews werden elektronisch an diesen Server gesandt. Hier werden die Interviews anonymisiert, so dass in der Auswertung nicht mehr ersichtlich ist, von wem welche Bewertung kam. Die Aufarbeitung von Interviewdaten erfolgt durch das *repertory grid*. Die mathematische Berechnung erfolgt durch Hauptkomponentenanalyse.

Innerhalb von 2-4 Minuten nach Versand des letzten Interviews gehen dem Anfordernden die Ergebnisse der Serverauswertung zu.

Die Auswertung erlaubt es unterschiedliche Perspektiven im Detail zu analysieren. So lassen sich beispielsweise Befragtengruppen getrennt auswerten.

Ausgewertete Daten werden sowohl in Excel®-Tabellen (siehe Abbildung 3) als auch in einer interaktiven grafischen Form zur Verfügung gestellt. Die Excel®-Tabelle arbeitet mit Farbunterlegungen, um zum Beispiel darauf hinzuweisen, ob die Situation des Vorhabens zum jeweiligen Zeitpunkt aus Sicht der angezeigten Gruppe positiv, zufriedenstellend oder kritisch bewertet wird.

Speziell geschulte Interpretatoren analysieren die Ergebnisse und erstellen einen etwa 8-seitigen Auswertungsbericht, der die zentralen Aussagen aller Interviews deutlich macht. Dabei wird sowohl auf Gemeinsamkeiten als auch auf spezifische Unterschiede zwischen den oder auch innerhalb der Interviewgruppen Bezug genommen.

Durch die Kategorisierung der gemachten Aussagen in verschiedene Themenblöcke kann neben der Analyse nach Erfolgsfaktoren bzw. nach Aktivitäten/Leistungen/Wirkungen zusätzlich eine Interpretation hinsichtlich der OECD/DAC-Kriterien, der Qualitätsprinzipien der GTZ und der Querschnittsthemen erfolgen.

ABBILDUNG 3: Beispiel für die Datenauswertung bei e-VAL®

Vergleichselement	Gruppe	Ausmaß, zu dem die angestrebte Situation zum aus- gewählten Zeitpunkt bereits erreicht ist				Anzahl Interviews
		Mittelwert	Streuung	Min	Max	
00.00. Situation bei Beginn der Ent- wicklungsmaßnahme: 2005	Alle Gruppen	14%	9%	0%	27%	10
00.02. Situation heute	Alle Gruppen	51%	17%	23%	77%	10
00.03. Situation in einem Jahr	Alle Gruppen	84%	7%	72%	100%	10
00.04. Situation nach Ende der Ent- wicklungsmaßnahme: 2012	Alle Gruppen	91%	9%	76%	100%	10
00.00. Situation bei Beginn der Ent- wicklungsmaßnahme: 2005	GTZ	21%	3%	16%	23%	3
00.02. Situation heute	GTZ	54%	13%	36%	64%	3
00.03. Situation in einem Jahr	GTZ	84%	12%	72%	100%	3
00.04. Situation nach Ende der Ent- wicklungsmaßnahme: 2012	GTZ	86%	10%	76%	100%	3
00.00. Situation bei Beginn der Ent- wicklungsmaßnahme: 2005	Mittler	7%	5%	0%	11%	3
00.02. Situation heute	Mittler	46%	12%	33%	62%	3
00.03. Situation in einem Jahr	Mittler	80%	3%	77%	84%	3
00.04. Situation nach Ende der Ent- wicklungsmaßnahme: 2012	Mittler	95%	7%	86%	100%	3
00.00. Situation bei Beginn der Ent- wicklungsmaßnahme: 2005	Partner	13%	10%	0%	27%	4
00.02. Situation heute	Partner	52%	22%	23%	77%	4
00.03. Situation in einem Jahr	Partner	87%	2%	85%	90%	4
00.04. Situation nach Ende der Ent- wicklungsmaßnahme: 2012	Partner	92%	8%	82%	100%	4

5. Nutzung der Ergebnisse

a) Vorhabenbezogene Nutzung

Die e-VAL[®]-Ergebnisse liefern insbesondere:

- Hinweise für TOR des Evaluationsteams;
- Inputs für Planung sowie Konzept- und Teamentwicklung;
- Daten für Bewertungen;
- Informationen für die Fortschrittsberichte der Vorhaben.

Die Ergebnisse von e-VAL[®] dienen in erster Linie dem Management der einzelnen Entwicklungsmaßnahmen zu Steuerungszwecken.

In Kombination mit den Ergebnissen des Monitoring des jeweiligen Vorhabens und gegebenenfalls weiteren quantitativen Erhebungen ist der e-VAL[®]-Bericht ein wesentlicher Input für Evaluierungen. Eingebettet in einen größeren Evaluierungs- oder Monitoringprozess liefern die e-VAL[®]-Ergebnisse zum einen Anhaltspunkte für die Formulierung weiterführender, spezifischer Fragestellungen. Zum anderen unterstützt der Interpretationsbericht von e-VAL[®] die Vorbereitung der Gutachter. Eine Vorstellung und Diskussion der e-VAL[®]-Ergebnisse stellt eine gute Einführung in die Evaluierungsarbeiten dar und erlaubt eine zusätzliche Fokussierung auf zentrale Aspekte des Vorhabens. Zusätzlich stellen die e-VAL[®]-Ergebnisse eine Datenquelle dar, die zum Abgleich der Evaluierungsergebnisse (Triangulation), für Schlussfolgerungen und Empfehlungen nützlich ist.

Seine Wirkungen kann e-VAL[®] dann in den Entwicklungsmaßnahmen entfalten, wenn aus den Ergebnissen angemessene Schlussfolgerungen gezogen werden:

- wenn z. B. „Strategie“ kritisch bewertet wird: Strategiediskussion;
- wenn z. B. „Zusammenarbeit nach Innen“ kritisch bewertet wird: Teambildung oder Input in Organisationsentwicklung;
- wenn z. B. die Erreichung von „Wirkungen“ kritisch bewertet wird: Analyse der Leistungserstellung des Vorhabens, Überprüfung von Wirkungshypothesen, Analyse der Angemessenheit von Zielen und Indikatoren;
- wenn z. B. die „Situation nach Ende des Vorhabens“ kritisch bewertet wird: Diskussion, wie die Nachhaltigkeit der Wirkungen sichergestellt werden kann.

b) Querschnittsanalyse

Durch die Möglichkeit der Kombination der e-VAL[®]-Ergebnisse verschiedener Vorhaben ergeben sich Möglichkeiten der Bewertung ausgewählter

Portfolios. Die Auswertung im Querschnitt kann nach Bedarf und nach unterschiedlichen Fragestellungen erfolgen. Fragestellung bzw. Kriterien können hierfür beispielsweise sein:

- Wie bewerten die Partner weltweit die Wirkung von GTZ-unterstützten Vorhaben auf *ownership*? Bestätigen die e-VAL[®]-Ergebnisse die Hypothese, dass die Wirksamkeit bezüglich *ownership* der Partner in Programmorientierter Gemeinschaftsfinanzierung (PGF)/Programme Based Approaches (PBA) besser als in alleinstehenden Projekten und Programmen ist?
- Welches sind aus Sicht der Befragten die Erfolgsfaktoren für die Wirksamkeit von Gesundheitsprogrammen in Afrika? Sind die Erfolgsfaktoren in Asien dieselben?
- Bewerten die Zielgruppen die Partizipation (als Voraussetzung für nachhaltige Wirksamkeit) in den Vorhaben in Lateinamerika genauso gut oder schlecht wie die GTZ-Befragten? Wie sieht das in Afrika aus?
- Wie wird in Asien der Beitrag der GTZ-unterstützten Vorhaben zur Erreichung übergeordneter Wirkungen (MDG) eingeschätzt? Gibt es Unterschiede zwischen Vietnam, Indien und den Philippinen?
- Die Bedeutung spezifischer Erfolgsfaktoren kann analysiert werden.
- OECD/DAC-Evaluierungskriterien können im Querschnitt über verschiedene (oder auch alle) Vorhaben evaluiert werden.

6. Anforderungen

Als Grundvoraussetzung für die Anwendung von e-VAL[®] müssen die Interviewer und Interpretatoren von Interviews über Computer, über die e-VAL[®]-Software, über einen Internetzugang und über die notwendigen Passwörter verfügen.

e-VAL[®] stellt hohe Ansprüche an Interviewer, die gut auf ihre Interviewpartner eingehen sowie flexibel und wortgewandt Fragen stellen können müssen, und an Interpretatoren von Interviews. Beide Personengruppen werden hierfür speziell ausgebildet. Nur Personal, das nach der Schulung für diese Tätigkeiten als qualifiziert eingeschätzt wird, darf Interviews durchführen und Auswertungsberichte schreiben. Dabei handelt es sich zum großen Teil um einheimische Fachkräfte, die zukünftig dieses Aufgabengebiet weitgehend vollständig abdecken sollen.

Ein von einem erfahrenen Interviewer geführtes Interview dauert 1,5 bis 2 Stunden. Der Gesamtaufwand für alle Interviews hängt von vielen

Faktoren ab. Falls die Interviews an einem zentralen Ort organisiert werden können, sind leicht 3 Personen pro Tag pro Interviewer zu befragen. Falls dies nicht möglich ist, sind zusätzlich Wege- und eventuell Reisezeiten zu berechnen, die unterschiedlich hoch ausfallen können.

Für die Interpretation der Ergebnisse aller Interviews einer Entwicklungsmaßnahme werden 2 Arbeitstage vorgesehen.

7. Besondere Eignung und Grenzen des Verfahrens

Eine wesentliche Stärke von e-VAL[®] liegt in der systematischen Erfassung, Darstellung und Nutzung der subjektiven Meinungen wichtiger Interessengruppen in einem Vorhaben. Dabei ermöglicht e-VAL[®] Einschätzungen zu Themen, die schwer „objektiv“ messbar sind, sich aber subjektiv oft treffsicher beurteilen lassen, z. B. *ownership* und *capacity development*.

Das Verfahren zeichnet sich auch durch seine Transparenz in der Durchführung der Interviews aus. Der Interviewablauf kann von den interviewten Personen gesteuert werden. Anhand der graphischen Darstellung und den ergänzenden Erklärungen der Interviewer können die Befragten ihr Interviewergebnis bestätigen oder korrigieren.

Das Verfahren ist allerdings ohne Zugang zu der spezifischen Software und zum e-VAL[®]-Server der GTZ nicht anwendbar. Zusätzlich ist eine eingehende Schulung in der Handhabung der Software unabdingbar. Das Verfahren benötigt spezifisch ausgebildetes Personal für die einzelnen Verfahrensschritte. Der für eine hohe Qualität notwendige Schulungs- und Backstopping-Bedarf ist erheblich.

Die Auswertungsergebnisse der Interviews liegen dem Auftragsverantwortlichen unmittelbar nach Versand der letzten Interviews in übersichtlicher Form vor. Sie können für die weitere Steuerung der Entwicklungsmaßnahme, zum Dialog mit den Partnern und für weitere Evaluierungs- und Monitoringarbeiten genutzt werden.

Die Ergebnisse von e-VAL[®] können für Analysen regional, sektoral oder für spezifische Organisationseinheiten aggregiert und nach unterschiedlichen Gesichtspunkten ausgewertet werden. e-VAL[®] erlaubt keine systematische Ausrichtung einzelner Interviews auf Wirkungen. Es liefert nur dann Daten zu Wirkungen, wenn die Interviewpartner diese von sich aus thematisieren. Allerdings können gemachte Aussagen zur Wirkungen in Querschnittsauswertungen systematisch ausgewertet werden.

e-VAL[®] liefert keine statistisch abgesicherten Daten, sondern macht Aussagen zu Tendenzen und Trends.

Da neben den Interviews keine weiteren Datenerhebungsmethoden vorgesehen sind, hängt die Qualität der Ergebnisse entscheidend von den

Detailkenntnissen der interviewten Person über die Entwicklungsmaßnahme und der Qualität der Interviewführung und Themenzuordnung durch die Interviewer ab. Die Qualität und Tiefe der Analyse des vorhabenspezifischen Interpretationsberichts entscheidet weitgehend über die Nützlichkeit und Verwertbarkeit von e-VAL[®]-Ergebnissen für Steuerungsprozesse und weitere Evaluierungsarbeiten.

Impact Monitoring and Assessment (IMA) nach Herweg/Steiner

BERNWARD CAUSEMANN / EBERHARD GOHL /
MATTHIAS LANZENDÖRFER

1. Allgemeine Informationen

Impact Monitoring and Assessment (IMA) wurde zuerst 1999 von Karl Herweg und Kurt Steiner veröffentlicht. Die Darstellung hier beruft sich auf die Fassung von 2002. IMA entstand aus einer Zusammenarbeit von GTZ und DEZA mit dem Centre for Development and Environment (CDE, Universität Bern) ab 1996 mit vielen internationalen Organisationen, darunter dem Rural Development Department der Weltbank, und wurde mit vielen Partnerorganisationen getestet.

Bei IMA handelt es sich nicht um ein „Verfahren“ im engeren Sinne, sondern um ein Konzept mit Bauanleitung (Teil 1 der IMA-Broschüre) und Instrumentenkasten (Teil 2). Es basiert auf einem didaktischen Herangehen, das Wirkungsmonitoring den beteiligten Projektmitarbeiter/-innen und den Zielgruppen vermitteln hilft.

Praktiker in der ländlichen Entwicklung und im Management natürlicher Ressourcen haben immer wieder auf IMA zurückgegriffen, u. a. im Rahmen von Trainingsmaßnahmen für die DSE, später InWEnt, in der Fortbildung von Partnerorganisationen von Misereor in Peru sowie bei Beratungseinsätzen zur Verbesserung des Wirkungsmonitoring von GTZ-geförderten Projekten.

2. Anwendungsbereiche

IMA wurde für Beratungsprojekte im Bereich des nachhaltigen Landmanagement entwickelt, ist aber leicht auf Beratungsvorhaben in ländlicher Entwicklung und Gemeinwesenentwicklung übertragbar. Das Verfahren ermöglicht die Analyse von Wirkungen solcher Entwicklungsmaßnahmen auf Mikro- und Mesoebene.

Die IMA-Instrumentensammlung ist für begleitende Beobachtung (Monitoring) entwickelt und unterstützt damit auch Evaluierungen. Sie

nimmt aber alle Phasen des Projektzyklus in den Blick und beinhaltet auch einen Verfahrensschritt zur Plananpassung. IMA erfasst sowohl Zustände als auch Prozesse, letztere vor allem im Sinne von Verläufen und Entwicklungen über die Zeit, wobei die Frage nach dem Wie der Wirkungserzielung zum Teil aufgenommen wird. Die Art der Datenerhebung, besonders die differenzierte Betrachtung der Stakeholder, hilft bei der Analyse der Prozesse.

3. Analyseansatz

Die Instrumentensammlung dient dazu, möglichst viele Wirkungen zu erfassen, geplante wie ungeplante, direkte und indirekte, leicht und schwer zuordenbare, auch negative, und dabei insbesondere auch Entwicklungen (nicht nur Wirkungen) im Kontext in den Blick zu bekommen.

Als Ansatzpunkt werden ausgehend von der Ursache (der 'Intervention') und dem Kontext die kombinierten Veränderungen betrachtet ('Bruttowirkung'). Dieses Konzept geht davon aus, dass die sogenannte „Zuordnungslücke“ nicht an einer bestimmten Stelle der Wirkungskette platziert ist, sondern in deren Verlauf immer größer wird. Vielfältige Ursachen werden systematisch erfasst und in Beziehung gesetzt, dabei werden subjektive und objektive Daten verbunden.

4. Durchführung

a) Verfahrensablauf

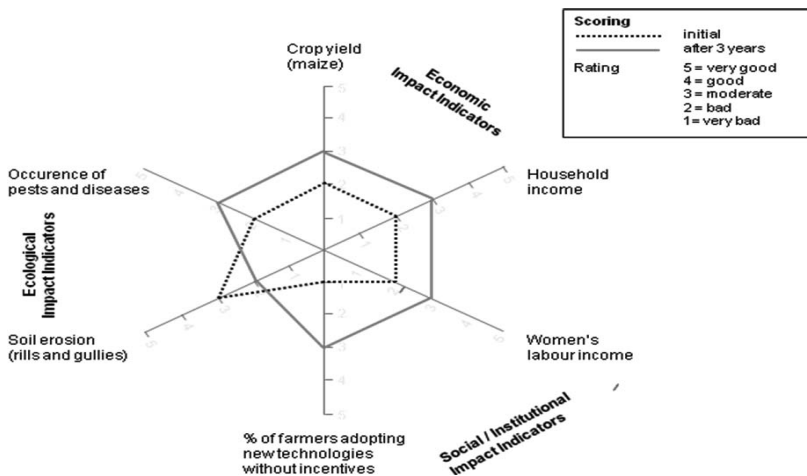
IMA gliedert sich in sechs Schritte:

1. Einbeziehung der Beteiligten und Informationsmanagement;
2. Review der Problemanalyse – vom Problembaum zum Flussdiagramm: eine Kontextanalyse mit nicht-linearen Einflüssen und Wechselwirkungen;
3. Formulierung von Wirkungshypothesen – und zwar sowohl positive wie negative;
4. Auswahl von Wirkungsindikatoren – diese sollen nicht bei der Planung eines Projekts, sondern in den ersten Monaten der Projektdurchführung entwickelt werden;
5. Entwicklung und Anwendung von Methoden der Wirkungsbeobachtung (Monitoring);
6. Wirkungseinschätzung (*assessment*).

IMA schlägt als zentrales Mittel ein Spinnendiagramm vor. Hier können Indikatoren anhand von Benchmarks analysiert, miteinander und über den Jahresverlauf verglichen werden. Methodische Anleitungen werden für folgende Aktivitäten gegeben:

- Beteiligtenanalyse mit Informationsplanung;
- partizipative Systemanalyse (Netzwerkanalyse), für die auch eine Excel-Vorlage existiert;
- Formulierung von spezifischen Wirkungshypothesen;
- Formulierung von spezifischen Wirkungsindikatoren;
- Interview, Diskussion, Foto-Monitoring, *transect walk* und Beobachtung;
- Erstellung von Spinnendiagramm (siehe Beispiel in Abbildung) und Wirkungsprofil.

ABBILDUNG 4: Beispiel für ein Spinnendiagramm in IMA nach Herweg/Steiner



b) Akteure

IMA ist als Monitoringinstrument konzipiert (das zur Selbstevaluierung hinführen kann) und wird daher im Wesentlichen von Projektleitung und -mitarbeitern durchgeführt, unter starker Beteiligung von verschiedenen anderen Stakeholdern. Externe Berater sind nicht explizit vorgesehen, können aber

einbezogen werden. Zielgruppen werden befragt und sollen eigene Einschätzungen abgeben. Stakeholdergruppen können ihr eigenes IMA durchführen; dies ist im vorliegenden Instrument jedoch nicht ausgearbeitet.

c) Datenerhebung

Um im Verlauf eines Projektes einen besseren Einblick in die Wirkungen und den Kontext zu bekommen, werden die Sichtweisen von Beteiligten und Zielgruppen und diverse objektive Daten in der Umwelt erhoben. Dabei ist keine systematische Stichprobenauswahl vorgesehen, sie kann aber je nach Bedarf nach Zufall oder aufgrund systematischer Kriterien geschehen.

Neben den unter 4a genannten Methoden werden im Rahmen von IMA Interviews (mit Fragebögen mit offenen und standardisierten Fragen) durchgeführt und, je nach Bedarf, verschiedene Instrumente des Participatory Rural Appraisal (PRA) genutzt.

Es sollten sowohl Indikatoren ausgewählt werden, die relativ kurzfristige Veränderungen zeigen, als auch solche, die erst auf längerfristige Veränderungen reagieren. Für diese braucht es zunächst eine *baseline*, dann können sie in unterschiedlichen zeitlichen Abständen erhoben werden. Indikatoren sollten bewusst so ausgewählt werden, dass sie für Aussagen über verschiedene Ebenen der Wirkungskette herangezogen werden können.

Verschiedene Perspektiven werden gezielt eingenommen und gesucht. Für alle Instrumente der Datenerhebung wird kurz beschrieben, wie sie zur Wahrnehmung verschiedener Perspektiven genutzt werden können.

d) Datenauswertung

IMA bietet verschiedene Instrumente für die Datenauswertung an, die je nach Bedarf angewandt werden sollen (z. B. Wirkungshypothesen, Spinnendiagramm, Wirkungsprofil, Netzwerkanalyse, Fragelisten zu Attribution und *follow-up*). Diese sind z. T. spezifisch nach Beteiligten oder entlang der Wirkungskette angelegt.

Jedes Projekt entwickelt sich sein eigenes Raster mit *benchmarks*. Dieses wird dann über mehrere Zeitpunkte angewandt und in einem Spinnendiagramm dargestellt. Der Umgang mit unterschiedlichen Perspektiven/Wahrnehmungen ist Teil der Auswertung.

Validiert werden die Ergebnisse durch eine Vielzahl von Instrumenten; die Instrumentensammlung lädt immer wieder ein zur Triangulation von Methoden und Quellen der Information.

5. Nutzung der Ergebnisse

Für Projektmitarbeiter bietet IMA eine hohe Transparenz; die Instrumente sind auf Verständlichkeit und Übersichtlichkeit angelegt. Die Transparenz der Wirkungsanalyseergebnisse für Externe und Zielgruppen hängt von der Ergebnisdarstellung ab. Die Publikationen legen viel Wert auf didaktische Vermittlung und Visualisierung.

Quantifizierung und Aggregation von Daten sind bei verschiedenen IMA-Instrumenten möglich, zentral im Spinnendiagramm. Es wird gezeigt, wie Messungen und Beschreibungen transparent in Skalenwerte übertragen werden können. Dies soll auch die Aggregierung erleichtern.

Um die Nutzung der Monitoringergebnisse zu fördern, ist die sechste Phase, Wirkungseinschätzung, darauf angelegt, zu Entscheidungen über Veränderungen im Programm zu führen. IMA ist vor allem für die interne Steuerung von Entwicklungsmaßnahmen gedacht. Ergebnisse können aber sehr gut auch nach außen zur Rechenschaftslegung und Öffentlichkeitsarbeit verwendet werden.

6. Anforderungen

Der Zeitaufwand für die Durchführung von IMA hängt davon ab, wie aufwändig das mit diesem Instrumentenkasten ausgestaltete Monitoringsystem ist. Die Instrumente können nach Bedarf ausgearbeitet und kombiniert werden. Die einzelnen Instrumente sind jedes für sich nicht sehr aufwändig, ein sinnvolles Monitoringsystem entsteht aber nur durch die Kombination mehrerer Instrumente.

Benötigt wird Kompetenz in der Anwendung von einfachen standardisierten und nicht-standardisierten Verfahren und bei der Datenanalyse sowie Vertrautheit mit Instrumenten des PRA.

7. Besondere Eignung und Grenzen des Verfahrens

Die IMA-Instrumentensammlung ist vor allem für das Monitoring von Wirkungen entwickelt worden und kann insofern wesentliche Daten für die Wirkungsanalyse liefern. Viele Elemente daraus können daher auch bei Evaluationen verwendet werden. Sinnvoll wäre es, eine Evaluation nach diesem Konzept frühzeitig zu planen, entsprechende Instrumente schon im Rahmen der Planung einzubeziehen und im Rahmen des Monitoring anzuwenden.

Der besondere Wert von IMA liegt in dem auf Partizipation angelegten Gesamtkonzept sowie in der hilfreichen (wenn auch nicht immer ästhetischen) Visualisierung der IMA-Broschüren. Aufgrund des didaktischen Auf-

baus eignen sie sich als Grundlage zur Wissensvermittlung über Wirkungsanalyse, auch wenn einige Begrifflichkeiten heute aktualisiert werden müssten.

Ein besonderer Aspekt von IMA ist der geschärfte Blick auf den Kontext des jeweiligen Projektes. IMA soll dazu dienen, sich auf einen veränderlichen Kontext einzustellen. Damit geht es davon aus, dass im Verlauf des Projektes auch größere Veränderungen der Planung notwendig werden.

Participatory Systems Analysis ist für die Systemanalyse besonders geeignet. Es liegt als Excel-Datei vor.

Eine Schwierigkeit besteht darin, dass IMA zwar so gestaltet ist, dass es so eng wie möglich an ein Project Cycle Management (PCM) angeschlossen werden kann, dass aber Organisationen häufig nicht bereit sind, zwei Systeme im Rahmen ihres PCM laufen zu lassen.

Die für den Schritt 6 (siehe 4a) vorgesehenen Instrumente zur Wirkungseinschätzung sind in der Broschüre sehr knapp dargestellt. Hier gibt es deutlich Bedarf an einer Weiterentwicklung des Verfahrens. Das Konzept geht davon aus, dass schon in der Planung und im Monitoring die Instrumente der Wirkungseinschätzung, insbesondere Indikatoren, überlegt und festgelegt werden.

Literaturhinweise

Herweg, Karl und Steiner, Kurt (2002): Impact Monitoring & Assessment. Instruments for use in rural development projects with a focus on sustainable land management. Volume 1: Procedure. Bern/Eschborn: CDE/GTZ et al. URL: <http://www.cde.unibe.ch/Tools/pdf/imavol1en.pdf> (4.5.2008).

Herweg, Karl und Steiner, Kurt (2002): Impact Monitoring & Assessment. Instruments for use in rural development projects with a focus on sustainable land management. Volume 2: Toolbox. Bern/Eschborn: CDE/GTZ et al. URL: <http://www.cde.unibe.ch/Tools/pdf/imavol2en.pdf> (4.5.2008). *Frühere Fassung:* Herweg, K.; Steiner, K. und Slats, J. (1999): Sustainable Land Management. Guidelines for Impact Monitoring. (4 Module). Bern/Eschborn: CDE/ GTZ et al. *Es gibt auch eine spanische Fassung unter* http://www.cde.unibe.ch/Tools/IM_Pub_Ts.asp

Herweg, Karl (2005): IMA – Main Messages and Experiences. Bern. (Link: www.impact-plus.de)

Graaff, Jan de (Hrsg.) (2007): Monitoring and evaluation of soil conservation and watershed development projects. Enfield, NH [etc.]: Science Publishers Publication.

Folgende Publikation gibt ein Beispiel eines Projekts, das in einem Workshop mit Participatory Systems Analysis arbeitet:

de Bruin, Annemarieke (2007): Managing a watershed by managing a project, 2005. Erosion Soil & Water Conservation Group, Soil Science Centre, Wageningen University. (Link:http://www.grupoice.com/esp/ele/manejo_cuencas/biblio/articulos/man_wat.html)

Impact Pathway

ACHIM ENGELHARDT

1. Allgemeine Informationen

Die Terminologie „Impact Pathway“ scheint keine eindeutige Quelle zu haben, sondern auf mindestens zwei unterschiedliche Quellen zurückzugehen: Douthwaite et al. (2002) beziehen sich in der Beschreibung des „Impact Pathway“ auf Roger et al. (2000) als einen Ansatz im Bereich der „Programme Theory Evaluation“. In diesem Zusammenhang wird der „Impact Pathway“ als eines der Instrumente beschrieben, das im Umfeld der Evaluierung von Sozialprogrammen entwickelt wurde, um Entscheidungen über Budgetverteilungen zu treffen. Bachmann (2006) hingegen sieht den Ursprung des „Impact Pathway“ jenseits der Evaluierung internationaler Entwicklungszusammenarbeit. In der Emissions-Modellierung und Evaluierung wird der Begriff „Impact Pathway“ erstmals in einer Serie von „Externen“-Projekten der Europäischen Kommission verwendet (Krewitt et al., 1998, 2001; Spadaro und Rabl, 1999).

Im Jahr 2003 wurde der „Impact Pathway“ von Engelhardt et al. als Methode empfohlen, um Wirkungen in der £ 250,000,000 „Renewable Natural Resources Research Strategy“ (RNRRS) des Britischen Ministeriums für Internationale Zusammenarbeit (DFID) zu analysieren, nachdem die Methode bereits erfolgreich in einem der zehn RNRRS⁵¹-Programme angewandt worden war.

2. Anwendungsbereiche

Der „Impact Pathway“ ist ein anwenderfreundliches Verfahren, um geplante Wirkungen im Kontext der internationalen Entwicklungszusammenarbeit in den unterschiedlichsten Sektoren visuell darzustellen. Die Methode ist in Bereichen der Forschung, Entwicklung der Privatwirtschaft, landwirtschaftliche Entwicklung und Öffentlichkeitsarbeit angewandt worden, doch steht sie prinzipiell allen Sektoren oder Interventionsarten offen. Der „Impact

⁵¹ Renewable Natural Resources Research Strategy

Pathway“ erlaubt Wirkungen auf verschiedenen Ebenen (Mikro-, Meso- und Makroebene) darzustellen und zu verfolgen, wobei sich die Mesoebene auf den institutionellen Rahmen eines Landes und die Makroebene auf die politische Ebene bezieht.

Der „Impact Pathway“ kann während verschiedener Phasen eines Interventionszyklus angewendet werden. Zu Beginn hilft das Verfahren Schlüssel-Akteure zu identifizieren, z. B. zusammen mit einer Stakeholder oder *forced field* Analyse. Zudem können das Management-Team der Intervention und die Partner ein gemeinsames Verständnis über Schlüssel-Aktoren, Institutionen und mögliche Veränderungsprozesse gewinnen. Die Identifizierung eines „Impact Pathway“ in der Planungsphase oft sehr komplexer Interventionen hilft, den Fokus auf Wirkungen zu behalten. Gleichzeitig richtet sich der Blick darauf, wie über die Intervention hinaus Wirkungen erreicht werden können.

Während des Interventionszyklus kann der „Impact Pathway“ dazu benutzt werden, ein Monitoring einsetzender Wirkungen zu betreiben. In der Mitte und/oder am Ende des Interventionszyklus können im „Impact Pathway“ dargestellte Wirkungen und deren Prozesse durch die dem „Impact Pathway“ zugrunde liegenden Daten intern oder extern evaluiert werden.

3. Analyseansatz

Um Wirkungen auf verschiedenen Interventionsebenen darzustellen, bietet sich der „Impact Pathway“ als ein mögliches Instrument an. Prozesse können aufgezeichnet werden, die auch zu unbeabsichtigten oder indirekten Wirkungen führen. Externe Faktoren, die eine Intervention oder deren Zielgruppen beeinflussen, sollten im Monitoring notiert werden, das die Anwendung des „Impact Pathway“ begleitet. Gleichzeitig kann so dokumentiert werden, wie und warum Veränderungen in den Koordinaten des „Impact Pathway“ zustande gekommen sind. Der „Impact Pathway“ ist ein explizites Modell für die Darstellung von Veränderungen durch eine Intervention und die Transformierung des Outputs der Intervention letztendlich bis zur *purpose*- oder Zielebene, z. B. institutionelle Reform oder Armutsminderung. Beim „Impact Pathway“ wird immer von *contribution* (dem Beitrag der Entwicklungsmaßnahme zu Zielen und Wirkungen), nicht *attribution* (Zuschreibung von Veränderungen zur Entwicklungsmaßnahme) ausgegangen. Jedoch bietet er die Möglichkeit, Prozesse aufzuzeigen, die gebündelt die „Zuordnungslücke“ verkleinern können.

4. Durchführung

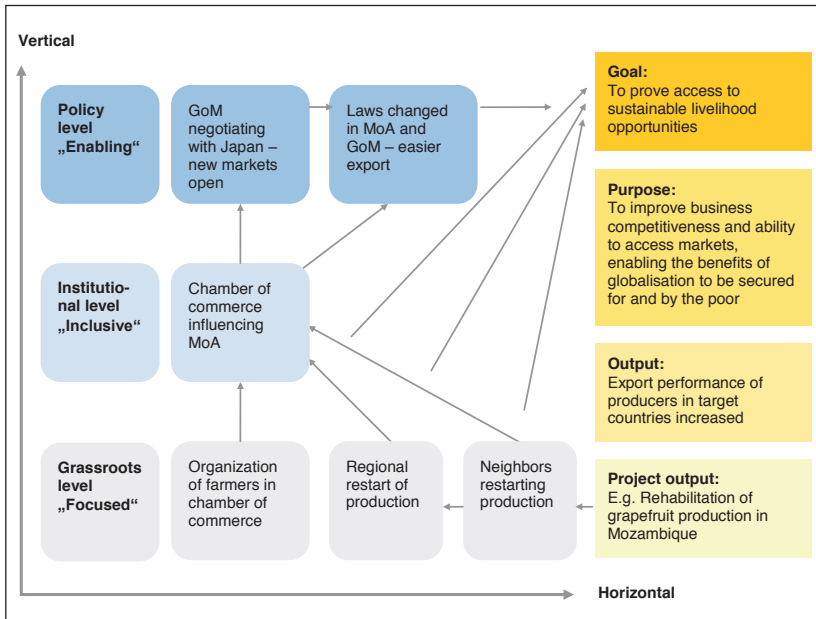
a) Verfahrensablauf

Die Anwendung des „Impact Pathway“ sollte auf partizipativen oder zumindest konsultativen Prinzipien beruhen. Die Triangulierung subjektiver Einschätzungen von Interventions-Managern, Partnern und anderen Schlüsselpersonen, dargestellt im „Impact Pathway“, aber auch externer Gruppen und der verschiedenen Zielgruppen durch Befragungen oder *story telling* tragen dazu bei, Wirkungen und deren Ursachen zu verifizieren. Die vereinfachte Darstellung eines „Impact Pathway“ (Abbildung 5) beruht auf einem Beispiel für die Stimulierung der Privatwirtschaft in Mosambik. Dazu wird die „grassroot“-Ebene, die institutionelle Ebene und die Politik-Ebene berücksichtigt. Wirkungen unterschiedlicher Art sind horizontal innerhalb einer Ebene oder vertikal zwischen Ebenen angedacht.

b) Akteure

Je nach verfügbarem Budget und Dauer der Intervention kann der „Impact Pathway“ entweder partizipativ oder konsultativ angewandt werden. Schlüsselpersonen haben die Gelegenheit, mögliche Veränderungsprozesse zu diskutieren und graphisch darzustellen, Monitoring durchzuführen und gegebenenfalls auch zu evaluieren. Dieser Prozess bietet besonders zu Beginn der Intervention die Möglichkeit, Vertrauen zwischen den unterschiedlichen Schlüsselpersonen zu schaffen und später gemeinsam während des Interventionszyklus zu lernen. Die Manager der Intervention, Partner und andere Schlüsselpersonen sollten während der gesamten Intervention an der Ausarbeitung der graphischen Darstellung von Veränderungsprozessen teilnehmen. Der Vorgang der graphischen Darstellung von Veränderungsprozessen auf verschiedenen Ebenen ist eine Möglichkeit, unterschiedliche Ansichten festzuhalten und zu diskutieren, um schließlich gemeinsam einen Kompromiss zu finden. Die visuelle Darstellung der Veränderungsprozesse stimuliert auch Monitoring-Treffen und ist für *mid-term*- oder finale Evaluierungen nützlich.

ABBILDUNG 5: „Impact Pathway“ am vereinfachten Beispiel der Entwicklung von Privatwirtschaft in Mosambik im Rahmen des DFID Business Linkages Challenge Fund



c) Datenerhebung

Nach der Erhebung von *baseline data* und einer *situation analysis*, können Veränderungsprozesse auf verschiedene Weisen aufgezeichnet werden. Ein solides Monitoringsystem ist für die Erhebung von Daten durch regelmäßiges *reporting* notwendig. Diese können durch Interviews mit Schlüsselpersonen oder Stichproben aus den Zielgruppen verifiziert werden. Dabei ist es wichtig, unterschiedliche Erfahrungen über die Art und Geschwindigkeit von Veränderungen festzuhalten. In regelmäßigen Abständen sollte der „Impact Pathway“ überprüft werden, um Veränderungen dank der Intervention aber auch des externen Umfeldes zu reflektieren. Neue Gelegenheiten für Veränderungsprozesse können somit identifiziert werden, ebenso wie Prozesse, die ins Stocken geraten sind.

Alle Veränderungen zum ursprünglichen „Impact Pathway“ aus der Planungsphase der Intervention sollten aufgezeichnet werden, um nachfolgende Evaluierungsaktivitäten zu erleichtern.

d) Datenauswertung

Bisher scheint die Anwendung des „Impact Pathway“ im Bereich der internationalen Entwicklungszusammenarbeit noch nicht direkt an IT-Komponenten zur Datenauswertung gekoppelt zu sein. Die Resultate des regelmäßigen Monitoring können in graphischer Form oder als Text in Berichte einfließen. Unterschiedliche Ansichten oder Meinungen können leicht zur Kenntnis genommen und in graphischen Darstellungen des „Impact Pathway“ berücksichtigt werden. Zur Verifizierung von Daten können Quellen trianguliert oder *spot checks* durchgeführt werden.

5. Nutzung der Ergebnisse

Die graphische Darstellung von Veränderungsprozessen durch den „Impact Pathway“ ist eine transparente Art und Weise, um über den Zeitraum einer Intervention Wirkungen zu dokumentieren und analysieren. Um diese Transparenz aufrecht zu erhalten ist es notwendig, Veränderungen im „Impact Pathway“ festzuhalten. Resultate sind qualitativer Art und können nicht aggregiert werden, doch erlaubt eine Ankopplung des „Impact Pathway“ an ein Monitoringsystem, z. B. mit Ampelsystem, die Quantifizierung einiger Resultate. Zudem können natürlich Schlüssel-Wirkungen über Modellierung errechnet werden.

Wie bei jedem M&E Verfahren erlauben Resultate des Monitoring, Korrekturen bei der Implementierung der Intervention durchzuführen, sie dienen der Rechenschaftslegung und Lernprozessen.

6. Anforderungen

Der „Impact Pathway“ ist unter Bedingungen zur Anwendung gekommen, in denen Standard-Verfahren für Planung, Monitoring und Evaluierung nur bedingt anwendbar waren, z. B. bei der Zusammenarbeit mit der Privatwirtschaft in *Public-Private Partnerships*. Neue Partner in der internationalen Entwicklungszusammenarbeit haben oft Schwierigkeiten den *logframe* korrekt anzuwenden und Zeitaufwand und Kosten für die Anwendung des „Impact Pathway“ erschienen angemessener. Mit anfänglicher Unterstützung einer M&E-Fachkraft können Interventionsleiter leicht in die Methodik des „Impact Pathway“ eingeführt werden. Zur Anwendung des Instruments ist nach einer kurzen Einführung kein spezielles Training notwendig, da der „Impact Pathway“ gerade von den Gruppen benutzt wird, die nicht mit zum Teil sehr arbeitsaufwändigen Methoden von M&E in der internationalen Entwicklungszusammenarbeit vertraut sind und daher leichter anwendbare Methoden bevorzugen.

7. Besondere Eignung und Grenzen des Verfahrens

Bei „Impact Pathway“ handelt es sich um ein besonders anwenderfreundliches Verfahren, das erfolgreich von neuen Partnern in der internationalen Entwicklungszusammenarbeit genutzt werden kann, die nicht an Standard-Verfahren der Rechenschaftslegung staatlicher Bürokratien gewöhnt sind. Die graphischen Darstellungen, auf denen das Verfahren basiert, sind einfach zu analysieren und zu interpretieren.

Wirkungen werde nicht mathematisch oder monetär dargestellt, sondern in Form von sozialen, politischen oder institutionellen Veränderungen, wobei der Fokus auf Kontribution, nicht Attribution liegt. Die Analyseergebnisse basieren auf qualitativen Daten, eine Aggregation der Resultate ist jedoch mittels der Anwendung eines begleitenden soliden Monitoringsystems teilweise möglich. „Impact Pathway“ kommt erst in Verbindung mit einem solchen Monitoringsystem zu seiner vollen Wirkung. Falls ein *logframe* oder andere Planungs- oder Wirkungsmodelle in der Interventionsplanung angewandt wurden, so können die *output/outcome*, *purpose* und *goal* Ebenen im „Impact Pathway“ übernommen werden.

Literaturhinweise

Bachmann, T. M. (2006): Hazardous substances and human health: Exposure, Impact and External Cost Assessment at the European Scale (Trace Metals and other Contaminants in the Environment).

- Briones, R.; Dey, M. M.; Ahmed, M.; Stobutzki, I.; Prein, M. und Acosta, B. O. (2004): Impact Pathway analysis for research planning: The case of aquatic resources research in the WorldFish Centre. In: NAGA, WorldFish Center Quarterly Vol.27 No.3 & 4 Jul-Dec 2004.
- Douthwaite, B.; Schulz, S.; Olanrewaju, A. und CGIAR (2002): Impact Pathway Evaluation: An Approach for Achieving and Attributing Impact in African Agriculture. www.ciat.cgiar.org/src/pdf/iita_bdouthwaite.pdf.
- Rogers, P. J.; Petersino, A.; Huebner, T.; und Hasci, T.A. (Hg.) (2000): Impact Pathway evaluation: practice, promise, and problems. In: Rogers, P. J.; Petersino, A.; Hasci, T.A. und Huebner, T. (Hg.): Impact Pathway in Evaluation. Challenges and Opportunities. New Directions for Evaluation 87.

Links zur Anwendung des Verfahrens sind zu finden unter:

- www.ciat.cgiar.org/src/pdf/iita_bdouthwaite.pdf.
- Engelhardt, A.; Pierce, J.; Bartram, E. and Ali, A. (DFID) (2003): Impact Assessment tools of DFID's Renewable Natural Resources Research Strategy. www.linux.parcinfo.org/index.php?option=com_docman&task=doc_download&gid=114&Itemid=146.

KAP-Studien

SIEGRID TAUTZ

1. Allgemeine Informationen

„KAP“ oder „KAB“ steht für die englische Bezeichnung *knowledge, attitude, practice (behaviour)* und hat sich auch in der deutschen Fachsprache unter der Bezeichnung eingebürgert. In der Entwicklungszusammenarbeit sind KAP-Studien eine erprobte Methode der Erhebung zielgruppenspezifischer Daten und eignen sich damit auch zur Wirkungsbeobachtung. KAP-Studien sind quantitative, repräsentative Erhebungen. KAP-Studien können immer dann angebracht sein, wenn zielgruppenspezifische Veränderungen in Wissen, Einstellung und Verhalten angestrebt werden, z. B. in der Gesundheitsaufklärung, der landwirtschaftlichen Beratung, der Gemeindebeteiligung, der Frauenförderung usw.

Im Folgenden wird an Beispielen aus dem Bereich Gesundheit die Bedeutung und Durchführung von KAP-Studien dargestellt.

Warum werden Informationen zu Wissen, Einstellungen und Verhalten erhoben? In frühen, konventionellen Ansätzen der Gesundheitserziehung und -beratung lag der Fokus auf Wissensvermittlung, in der Annahme, dass daraus automatisch entsprechende Verhaltensänderungen erfolgen würden. Jahrzehntelange Erfahrungen aus der Ernährungsberatung, Raucheraufklärung u. a. haben auch in Europa gezeigt, dass Verhalten(sänderung) neben ausreichender und angemessener Information von vielen Faktoren beeinflusst wird. So sind etwa im Bereich Sexualität und reproduktives Verhalten von vorherrschenden Normen geprägte persönliche Einstellungen für die Übersetzung von Kenntnissen in entsprechendes Verhalten entscheidend. Dies erfordert entsprechende Erhebungsmethoden.

2. Anwendungsbereiche

KAP-Erhebungen werden nicht ausschließlich, jedoch besonders häufig im Gesundheitsbereich im weiteren Sinn, überall dort wo es um die Veränderung von Wissen, Einstellungen und Verhalten geht, angewandt. Dies kann die Änderung von Ernährungsgewohnheiten oder Hygieneverhalten (Benut-

zung von Toiletten/Latrinen, geschützten Wasserquellen, u. a.) sein. Historisch haben sie besonders mit der Einführung der umfangreichen Familienplanungsprogramme der 70-er und 80-er Jahre an Bedeutung gewonnen; heute ist dies der Bereich „Sexuelle, reproduktive Gesundheit und Rechte“ (SRHR). In vielen Maßnahmen der deutschen EZ, sowohl der technischen Zusammenarbeit (GTZ), als auch z. B. der Social Marketing-Maßnahmen der Finanziellen Zusammenarbeit (FZ, KfW) oder der Aufklärungsprojekte von NRO wie der Deutschen Stiftung Weltbevölkerung, sind KAP-Studien wichtige Methoden zur Wirkungsbeobachtung. Prinzipiell kann auf allen Ebenen (Mikro/Meso/Makro) mit Hilfe von KAP analysiert werden so lange eine ausreichende Stichprobe für aussagekräftige, repräsentative Ergebnisse genommen wird.

Zur Beurteilung der Wirkungen von Maßnahmen ist erforderlich, dass Ausgangsdaten (*baseline*) erhoben werden und in angemessenem Abstand (2-3 Jahre) eine Wiederholungsstudie durchgeführt wird. Leider wird Ersteres häufig vergessen und erst kurz vor Ende einer Intervention oder im Rahmen einer Evaluierung daran gedacht, „mal zu schauen, was die Maßnahme gebracht hat“.

3. Analyseansatz

Im Wesentlichen bildet eine KAP-Studie einen punktuellen Querschnitt (Wissen, Einstellungen, aktuelles Verhalten) ab. Wiederholungsstudien können die Veränderung von Zuständen erfassen. Dies erlaubt je nach Fragestellungen auch die Beurteilung von Prozessen. KAP-Analysen konzentrieren sich vor allem auf geplante, weniger auf ungeplante Wirkungen. Sie sind je nach Formulierung der Fragen und Auswahl der Befragten geeignet, sowohl direkte als auch indirekte Wirkungen zu erfassen. KAP-Analysen erfassen die subjektive Einschätzung der Befragten. Es ist möglich, Fragen zum Zuordnungsproblem, d. h. zur Sicht der Befragten auf den Zusammenhang von Maßnahmen und Wirkungen, zu stellen, dies ist jedoch nicht Kern der KAP-Studien.

4. Durchführung

a) Verfahrensablauf

Der Zeitpunkt für die Erhebungen wird abhängig vom Bedarf an Ausgangsdaten sowie von Daten zur Bewertung von Wirkungen einer Maßnahme festgelegt. Fragebögen werden nach den unten dargestellten Krite-

rien entwickelt, getestet und bei den ausgewählten Befragten angewandt. Die Ergebnisse werden analysiert und ggf. wird in einer folgende Befragung der Fragebogen angepasst.

b) Akteure

Auswahl der Teilnehmenden

Die Auswahl der Studienteilnehmer muss, um Kriterien für Repräsentativität zu genügen, nach dem Zufallsprinzip erfolgen. Da es sich im Rahmen der EZ jedoch i. d. R. um anwendungsorientierte Forschung handelt, die oft unter logistisch erschwerten Bedingungen (z. B. unzuverlässige Melderegister) und mit sensiblen Themen in soziokulturell unterschiedlich offenen Kontexten stattfindet, müssen zuweilen pragmatische Kompromisse eingegangen werden. So wurde etwa bei einer Studie im Rahmen eines GTZ-unterstützten Vorhabens in Marokko ein für fundamentalistische Orientierung bekannter Schulbezirk aus der Stichprobenwahl ausgeschlossen, um nicht die Gesamtdurchführung zu gefährden.

Ob die Zufalls-Auswahl in Stufen erfolgt oder aus einer Liste aller Schulen der Region randomisiert ausgewählt wird, und ob stratifiziert wird, etwa nach Geschlecht (so dass unabhängig von der tatsächlichen Proportion der Anteil weiblicher und männlicher Befragter etwa gleich ist), hängt von Fragestellung und Kontext ab.

Da die individuellen Subjekte für die Wiederholungsstudie ja in der Regel aus verschiedenen Gründen (Migration, natürliches Herauswachsen aus der relevanten Altersgruppe u. a.) nicht identisch mit denen der Ausgangsstudie sein können, müssen zur Sicherstellung der Vergleichbarkeit der Ergebnisse äußerst sorgfältig Kriterien der Teilnehmerauswahl definiert werden. Während die Kernfragen identisch sein müssen, können für die Folgestudie(n) durchaus Fragen, die den Prozess und Kontext der Maßnahme reflektieren (z. B. Nutzung der Leistungen; bei Jugendlichen Reaktionen der Familie und sozialen Umgebung), hinzugefügt werden, um auch unbeabsichtigte und über die direkte Nutzung von Outputs hinausgehende Wirkungen zu erfassen.

Einbeziehung weiterer Akteure

Wichtige Voraussetzungen für die reibungslose Durchführung von KAP-Studien, besonders wenn junge, unverheiratete Menschen beteiligt sind, sind im Vorfeld die Sensibilisierung, Information und gegebenenfalls Einbeziehung von Entscheidungsträgern und wichtigen Akteuren (Schulbehörden, Gemeindegremien, Elterngremien, religiöse Führer, Organisationen der

Zivilgesellschaft u. a.). Die Erhebung kann mit externer Unterstützung durchgeführt werden, die erfahrungsgemäß jedoch im Hintergrund bleiben sollte. Sowohl für die Akzeptanz einer solchen Studie seitens der Bevölkerung als auch für die Aneignung der Ergebnisse ist eine möglichst weitgehende Durchführung durch intensiv angeleitete lokale Fachkräfte und Interviewer wünschenswert.

c) Datenerhebung

Prinzipiell gibt es drei Möglichkeiten der Anwendung: elektronische oder papiergebundene Befragung sowie persönliche Interviews.

Die elektronische Befragung via Internet spielt in der EZ noch keine Rolle. Zunehmend kommen aber auch in armen Ländern Eingabecomputer zur Anwendung, in die der Befragende die Antworten eintippt oder der Befragte am Computer selbst eingibt. Dies Verfahren verringert die Fehlerwahrscheinlichkeit bei der Dateneingabe und erspart den ersten Schritt der Datenverarbeitung (Eingabe).

Ob man sich für das individuelle Ausfüllen der Fragebögen durch die Befragten entscheidet oder die Fragebögen nach genau zu befolgenden Standards durch Interviewer ausfüllen lässt, hängt vom Bildungs- und Alphabetisierungsstand der Zielgruppe sowie ihrer Kompetenz im Umgang mit derartigen Instrumenten ab. Für die Vermeidung von systematischen Fehlern (z. B. Höflichkeitsbias der Befragten) und der Sicherstellung größtmöglicher Validität ist hier ein äußerst sorgfältiges Interviewer-Training erforderlich.

Unerlässlich ist ein Pre-Test des Instruments mit einer der Gruppe von Befragten, die die Charakteristika der Zielgruppe möglichst genau repräsentiert. Dies trägt wesentlich zur Relevanz des Instruments, Vermeidung von Fehlerquellen und damit zur Validität der Daten bei.

d) Frageformen und Antwortmöglichkeiten

Die Frageinstrumente dürfen v. a. in Kontexten, die mit diesem Medium nicht vertraut sind, keineswegs zu lang sein, d. h. 30-40 Fragen nicht überschreiten. Eine günstige „Dramaturgie“ ist, die Befragung mit einfachen, interessanten Fragen zu beginnen, eventuell tabubehaftete, schwierigere Fragen in den Mittelteil zu stellen und die Befragung mit schnell zu beantwortenden Fragen zur Sozialstatistik/Demographie abzuschließen. Folgende Fragekategorien bieten sich an. In Klammern steht jeweils ein Hinweis darauf, ob sie sich vorwiegend auf Wissen, Einstellung oder Verhalten beziehen.

Ja/nein-Fragen (geschlossene Fragen)

Ja/nein-Fragen lassen nur zwei Antwortmöglichkeiten zu. Es ist ratsam, zusätzlich „weiß nicht“ als Antwortmöglichkeit anzubieten.

Beispiel:

- Sollte die Abgabe von Verhütungsmitteln an Minderjährige erlaubt sein? (Einstellung) – Ja / nein / weiß nicht
- Haben Sie beim letzten Geschlechtsverkehr ein Kondom benutzt? (Verhalten) – Ja / nein / weiß nicht

Multiple Choice

Dafür eignen sich vor allem Wissensfragen. Es besteht die Möglichkeit Antwortoptionen, die nur eine einzig richtige Antwort oder mehrere richtige Antworten zulassen, anzubieten. Dies sollte angegeben werden („mehrere Antworten sind erlaubt“; „bitte entscheiden sie sich für eine Antwort“).

Beispiel:

Wie kann HIV übertragen werden? (Wissen)

- Wenn man aus derselben Flasche trinkt wie ein HIV-Infizierter
- Wenn man ungeschützten Geschlechtsverkehr mit einem HIV-Infizierten hat
- Wenn man einen AIDS-Kranken umarmt
- Durch den Kontakt mit Instrumenten, die mit dem Blut eines HIV-Infizierten verunreinigt sind (beim Friseur, Zahnarzt, bei Spritzen, z. B. bei Drogenabhängigen)
- Durch die Benutzung von Toiletten, Duschen oder Badewannen, die auch HIV-Infizierte benutzen
- Durch Insekten, die zuvor einen Infizierten gestochen haben
- Von einer infizierten Mutter auf ihr ungeborenes Kind
- Ich weiß nicht

Offene Fragen

Diese sollte man in einer KAP-Studie möglichst vermeiden. In der Vorbereitung kann man durch einen Pre-Test in aller Regel angemessene Antwortoptionen herausfinden. So könnte eine Meinungsäußerung zu Fragen wie „Was halten Sie von Empfängnisverhütung?“ (Einstellung) oder „Wie beurteilst Du den Erfolg der Aufklärungsmaßnahmen?“ (Frage zum Projekt/Beurteilung) in Form von geschlossenen Optionen erfragt werden.

Statements

Sie eignen sich besonders zur Untersuchung von Einstellungen, da die Befragten aufgefordert werden zuzustimmen oder abzulehnen.

Beispiel:

- Verhütung ist Frauensache.
- AIDS-Kranke sollten keine öffentlichen Schulen besuchen dürfen.

Stimme zu / stimme nicht zu / bin unentschieden; oder skalierte Antwort

Eingruppierungs-Fragen

Bei Eingruppierungsfragen werden Bereiche oder Gruppen vorgegeben, in die sich der/die Befragte einordnen muss.

Beispiel:

Sind Sie: heterosexuell / homosexuell / bisexuell

Eine Gruppierung der Daten im Prozess der Analyse ist oft sinnvoller. Zum einen wird die Eingabe durch Gruppierung nicht erleichtert. Außerdem beschneidet man unnötig Auswertungsmöglichkeiten wie die Berechnung des Median und Mittelwertes.

Skalierte Fragen

Beispiele mit einer Skala von z. B. 1 bis 6 (hier ist zu beachten, dass die Wahrnehmung von niedriger bzw. hoher Ziffer als gut oder schlecht kulturspezifisch ist):

- Wie gut gefallen Dir die Cartoons und Abbildungen in den HIV-Informationsbroschüren? (Frage zum Projekt/Beurteilung; Einstellung)
Sehr gut (1).....gar nicht (6)
- Schlafen Ihre Kinder unter einem Moskitonetz? (Verhalten)
Immer / meist / gelegentlich / nie / weiß nicht

Rangfrage

Beispiel (Einstellung):

Ordne folgende Antworten nach „erstens zweitens drittens“ etc.:

Am wichtigsten ist mir:

Schule / Gesundheit / Partnerschaft / Freunde / Spaß / Liebe / Sex / Abenteuer

Gerade oder ungerade Zahl von Antwortoptionen?

Bei Skalen mit einer ungeraden Zahl (z. B. 1 bis 3 oder 5) neigen manche Befragte aus unterschiedlichen Gründen dazu, den mittleren Wert zu wählen mit entsprechenden Folgen für ein valides Auswertungsergebnis. Daher wird eine mittlere Kategorie häufig bewusst weggelassen um einen klaren – positiven oder negativen – Trend erkennen zu können.

e) Datenauswertung und -interpretation

Für die Datenauswertung werden die für standardisierte Erhebungen gängigen Instrumente und Programme verwendet, also z. B. EPI Info oder SPSS. Die Antworten auf offene Fragen müssen kodiert und dann eingegeben und ausgewertet werden.

Zuverlässigkeit und Qualität der Ergebnisse

Befragungen von Menschen können nur dann die Wirklichkeit zutreffend beschreiben, wenn die Befragten die Antworten verstehen und wahrheitsgemäß beantworten.

Fragen können z. B. dann nicht verstanden werden, wenn sie auf einem Weltbild beruhen, das dem Befragten fremd ist. Die einfache Frage „wie viele Kinder möchtest Du haben“ setzt implizit voraus, dass Frau/Mann glaubt, das beeinflussen zu können. Eine häufige Antwort in KAP-Studien zur Familienplanung in Afrika war, „So viel mir Gott schenkt“.

Bei der Analyse der Ergebnisse muss bedacht werden, dass das „P“ oder „B“ in der Abkürzung der KAP-Studie für berichtetes Verhalten steht. Ob z. B. tatsächlich 60% der befragten Prostituierten bei jedem Sexualkontakt Kondome benutzen, könnte nur mit einer anderen Untersuchungsmethode herausgefunden werden, z. B. mittels simulierter Klienten. Besonders bei einem engagierten Programm zur Kondomnutzung kann es durchaus sein, dass der soziale Druck, das gewünschte Verhalten zu berichten, erhöht wurde (*social desirability bias*).

Bei Studien zur Jugendsexualität haben sich in manchen kulturellen Kontexten bzgl. Fragen des sexuellen Debuts und der Häufigkeit sexueller Kontakte ein Über-Berichten der männlichen und Unter-Berichten der weiblichen Befragten gezeigt. Diese Gefahr kann durch geschickt gestellte Kontrollfragen im Fragebogen etwas reduziert oder durch ergänzende qualitative Erhebungen korrigiert werden.

Wichtig ist auch eine realistische Bewertung der Ergebnisse insofern, als keine Maßnahme Erfolge in Bezug auf Kenntnisse, Einstellungen und v. a. Verhalten für sich alleine reklamieren kann, die Zuordnungslücke hier

also besonders deutlich sein kann. Bevölkerungsgruppen leben nicht in einem Vakuum und sind nicht nur unterschiedlichen Informationsquellen sondern auch anderen etwaigen Einflussfaktoren ausgesetzt, wie z. B. Migration, fallendem/steigendem Einkommen, besserem Zugang zu Gesundheitsdiensten, etc. Ein Zusammenhang zwischen Maßnahme und – positiven, negativen, beabsichtigten und unbeabsichtigten – Wirkungen kann, wie insgesamt im Rahmen der Wirkungsbeobachtung auch hier nur plausibel hergestellt werden, umso mehr in Kontexten mit vielen Akteuren, z. B. sektorweiten Ansätzen (Soziale Wirkungsanalyse in armutsorientierten Projekten (SWAP)).

Die Validierung erfolgt durch Kontrollfragen und Triangulation, z. B. Ergänzung durch semi- oder unstrukturierte Interviews.

5. Nutzung der Ergebnisse

Die Ergebnisse sind weitgehend quantitativ und damit aggregierbar. Sie sind nachvollziehbar und gut verständlich.

Nutzer der Ergebnisse sind zunächst meist die Programmbeteiligten auf lokaler, regionaler und nationaler Ebene, für die sie als Grundlage des Monitoring und eventuell erforderlicher Anpassungen der Maßnahme bedeutsam sind. Häufig werden die Ergebnisse jedoch für eine Reflexion im weiteren Rahmen genutzt und eventuell publiziert. In Marokko (GTZ) entwickelte sich daraus z. B. eine regionale Jugendkonferenz, auf der Erfahrungen aus der ganzen Region Mittlerer Osten/Nordafrika diskutiert und entsprechende Lehren abgeleitet wurden.

6. Anforderungen

Der Aufwand für die Durchführung kann in Abhängigkeit von der Größe der Stichprobe und örtlichen Gegebenheiten (Infrastruktur, Einzugsgebiet, Vorhandensein spezifischer Kompetenz zur Durchführung und statistischen Auswertung u. a.) erheblich sein. Er ist jedoch erfahrungsgemäß eine sinnvolle Investition, da die darauf aufbauenden Maßnahmen dadurch oft den „Segen“ der Evidenzbasierung haben, was gerade in traditionell sensiblen Themenbereichen wichtig sein kann. Unerlässlich ist die Einhaltung ethischer Standards. Dazu gehört, dass Interviewpartner über die Ziele der Datenerhebung und Verwendung der Ergebnisse informiert werden müssen und nur Personen befragt werden, die ausdrücklich ihre Zustimmung geben. Daten müssen strikt vertraulich behandelt werden. In vielen Ländern müs-

sen Studien-Vorschläge einem Ethikkomitee zur Begutachtung vorgelegt werden.

KAP-Studien sollten mindestens zu zwei verschiedenen Zeitpunkten durchgeführt werden, um eine Vergleichsbasis zu haben und Wirkungen feststellen zu können. Sie können auch in ein Monitoringsystem integriert werden.

7. Grenzen und Herausforderungen des Verfahrens

Die Ergebnisse von KAP-Studien können nur so aussagekräftig sein wie die gestellten Fragen dies zulassen. Dies wird leider viel zu wenig beachtet. So werden weltweit Standardfragebögen zu sexueller und reproduktiver Gesundheit eingesetzt, ein Bereich, der ganz besonders von kulturspezifischen Aspekten beeinflusst ist. Wenn jedoch potenziell wichtige Fragen nicht gestellt werden oder die gestellten Fragen nur begrenzt relevant sind, dann spiegeln auch die Ergebnisse ein zum Teil verzerrtes und unstimmiges Bild der Realität der Zielgruppe. Dies wiederum ist keine solide Basis für die Planung und die Wirkungsbeobachtung einer Intervention.

Daher ist in Kontexten, über die noch wenig verlässliche Vorinformation über die Zielgruppe besteht, eine qualitative Vorstudie erforderlich, aus deren Ergebnissen dann relevante Fragestellungen generiert werden können. Umgekehrt ist es oft sinnvoll, quantitative Erhebungen wie KAP-Studien durch qualitative semi- oder unstrukturierte Interviews (Gruppen- oder individuelle) zu ergänzen (Datentriangulierung) und/oder zu vertiefen.

Literaturhinweise

Schnell, R.; Hill, P.B. und Esser, E. (2008): Methoden der empirischen Sozialforschung. München: Oldenbourg.

World Health Organization, International Development Research Centre, (2003): Designing and conducting health systems research projects, Volume 1: Proposal development and fieldwork; Volume 2: Data analysis and report writing. Amsterdam: KIT Publishers.

Beide Bände zum Herunterladen:

[http://www.kit.nl/smartsite.shtml?ch= FAB&id =83 32&ItemID=1091&RecordTitle= Designing%20and%20conducting%20health% 20system%20research%20projects](http://www.kit.nl/smartsite.shtml?ch=FAB&id=8332&ItemID=1091&RecordTitle=Designing%20and%20conducting%20health%20system%20research%20projects) (29.1.09).

Method for Impact Assessment of Programmes and Projects (MAPP)

SUSANNE NEUBERT / REGINA MÜLLER

1. Allgemeine Informationen

MAPP ist ein partizipatives Verfahren der Wirkungsanalyse, mit dessen Hilfe die Nutzer- und Mittlersicht systematisch erhoben wird. Auf Grundlage von Gruppendiskussionen werden in einer festen Abfolge sechs bis acht aufeinander aufbauende Instrumente eingesetzt, die zusammengekommen eine aussagekräftige Einschätzung der Veränderungen vor Ort, der Zuordnung der durchgeführten Entwicklungsmaßnahmen sowie der intendierten und nicht intendierten Wirkungen erlauben. Die Methodik wurde 1999 am Deutschen Institut für Entwicklungspolitik (DIE) entwickelt, zunächst hieß sie „Soziale Wirkungsanalyse in armutsorientierten Projekten“ (SWAP), seit 2000 trägt sie die Bezeichnung MAPP.

Sowohl staatliche als auch nicht-staatliche EZ-Organisationen bedienen sich des Verfahrens, außerdem wurde MAPP bereits von mehreren Consulting-Firmen und freien Gutachtern angewandt. Konkret wurde das Verfahren zur Evaluierung von GTZ-, KfW-, Weltbank-, DEZA- und nationalen Programmen der Partnerländer verwendet. Auch setzen es lokale Initiativen und Organisationen der kirchlichen EZ ein.

2. Anwendungsbereiche

Mit MAPP kann eine große Bandbreite von Programmtypen evaluiert werden: Erprobte Sektoren sind Ressourcenmanagement, Landwirtschaft und ländliche Entwicklung, Wasser, Fischerei, Dezentralisierung, Gesundheit und Bildung. MAPP prüft de facto Wirkungen und setzt damit voraus, dass eine gewisse Projektdauer bereits verstrichen ist, so dass Wirkungen überhaupt eingetreten sein können. Die Wirkungen werden retrospektiv ermittelt, d. h. eine *baseline study* muss nicht vorhanden sein.

MAPP kann auch bei Sektor- oder Portfolioanalysen herangezogen werden, zweckmäßig ist dann eine Kombination mit weiteren Verfahren, wie dem Rapid Sectoral and Rapid Portfolio Assessment (RSA) (siehe Ab-

schnitt 2.3. in diesem Handbuch). Das Verfahren ist geeignet für jede Art von Entwicklungsmaßnahmen der finanziellen und technischen Zusammenarbeit, für lokale bis flächendeckende Programme, die letztlich auf die Bevölkerungsebene abzielen. Prinzipiell können alle Ebenen (Mikro-, Meso- und Makroebene) mit Hilfe von MAPP analysiert werden, es kommt auf die Anzahl der Stichproben an sowie auf die befragten Stakeholdergruppen, ob ein *scaling up* möglich ist. Prinzipiell können die Ergebnisse (Punktwerte) aus einzelnen Durchgängen aggregiert werden. Mit MAPP wurden bereits Erfahrungen in allen Kontinenten gesammelt.

3. Analyseansatz

Das Verfahren ist prozessorientiert, d. h. es werden retrospektiv mehrere Zeitpunkte über den gesamten Zeitraum des Programms/Projekts erfasst und jeweils Punktwerte entlang von Kriterien erteilt. Dabei kommt es weniger auf die absoluten Punkte an als vielmehr auf die Veränderungen der Punktzahlen über die Zeit, wobei die jeweiligen Ursachen dann qualitativ begründet werden.

MAPP erfasst im ersten Schritt die Veränderungen des Lebenskontextes der Menschen über die Zeit (z. B. zehn Zeitpunkte) und ordnet erst im zweiten Schritt die Ursachen für diese Veränderungen zu (externe Rahmenbedingungen oder Entwicklungsmaßnahmen). Es werden dabei auch Programme anderer Geber in der Zielregion in die Betrachtung einbezogen. MAPP erfasst den *impact* und *outcome* von Maßnahmen, einschließlich geplanter und ungeplanter, direkter und indirekter, positiver oder negativer Wirkungen.

Die „Zuordnungslücke“ wird bei MAPP mit einer Einflussmatrix überbrückt, bei der Einflussstärken der Maßnahmen zugeordnet, gewichtet und begründet werden. Alle Maßnahmen werden im Hinblick auf ihre Wirkung auf alle Kriterien evaluiert, die zusammengekommen das Zielkonzept beschreiben. Die im Diskurs gewonnenen Ergebnisse werden zunächst kommunikativ und dann mit Hilfe der Triangulation validiert (Mehraugenprinzip, *cross-checking* mit Daten aus anderen Quellen und Dokumenten). Insgesamt beruht die Wirkungseinschätzung auf subjektiven Einschätzungen / Wahrnehmungen, und in der Praxis werden vornehmlich Zielgruppen befragt. Die Diskussionsgruppen sollen allerdings heterogen zusammengesetzt sein, so dass auch gar nicht oder negativ Betroffene Stakeholder daran teilnehmen (z. B. besonders Arme, Frauen etc.). Prinzipiell spricht nichts gegen die zusätzliche Befragung von Kontrollgruppen. Dies wäre mit dem gleichen Verfahren möglich.

4. Durchführung

a) Verfahrensablauf

Bei MAPP werden 6 bis 8 Instrumente in einer logischen Abfolge eingesetzt. Je nach Erkenntniszweck kann sich die Untersuchung auch nur auf einen Teil der Instrumente stützen. Die Minimalvariante von MAPP besteht aus der Anwendung der (1) Lebenslinie, (2) Trendanalyse, (5) Einflussmatrix und dem (6) Entwicklungs- und Wirkungsprofil.

1. Lebenslinie

Sie zeigt die Gesamtentwicklung in der Projektregion über einen Zeitraum, der die Laufzeit des zu evaluierenden Programms oder Projekts knapp überschreitet. Die Kurvenausschläge zeigen die Veränderungen in positiver oder negativer Richtung entlang einer Fünf-Punkte-Skala. Externe Ereignisse, Krisen und Schocks kommen hierbei ebenso zur Sprache wie die wichtigsten Ereignisse im Hinblick auf den Beginn oder den Abschluss von Entwicklungsmaßnahmen oder Eigeninitiativen.

2. Trendanalyse

Sie erfasst die Veränderungen im Hinblick auf vorher definierte Kriterien, die gemeinsam das Zielkonzept beschreiben (z. B. Armutsminderung). Beispielsweise werden 10 Kriterien für Armut definiert und jeweils den Schlüsselprozessen von Armutsminderung zugeordnet. Diese sind z. B. der Lebensstandard, der Zugang zu Ressourcen, die Erweiterung des Wissens sowie die Partizipation an Rechten und Macht. Der Betrachtungszeitraum startet vor Beginn der zu evaluierenden Entwicklungsmaßnahme und endet in der Gegenwart. Die Bewertungen erfolgen nach einer Punkteskala von „sehr schlecht“ bis „sehr gut“. Abschließend wird der Gesamttrend für jedes Kriterium ermittelt.

3. Cross-checking

Dieses ist sowohl für die Überprüfung der Informationen über die Programmumgebung als auch über die Maßnahmen nützlich. Neben dem Mehraugenprinzip, das bei MAPP durchgehend Anwendung findet, werden Statistiken, Beobachtungen etc. herangezogen, um die Daten zu kontrollieren.

4. Maßnahmenliste

Hier wird die Relevanz der Maßnahmen aus Sicht der Nutzerinnen und Nutzer (Zielgruppen) und die Einordnung in ihren Arbeits- und Finanzkontext

vorgenommen. Aus den Informationen können Schlussfolgerungen zu den Maßnahmen aller aktiven Geberorganisationen, zur Signifikanz, Relevanz und zur Chance auf Nachhaltigkeit gezogen werden

5. Einflussmatrix (Abbildung 6)

Mit ihrer Hilfe erfolgt die Wirkungszuschreibung von Maßnahmen auf die Entwicklung eines jeden Kriteriums („Nettowirkung“) und schließlich auf das gesamte Zielkonzept. Anschließend werden Passiv- und Aktivsummen gebildet. Die Aktivsummen zeigen, welche Maßnahmen Schlüsselwirkungen auf die Armutskriterien haben, die Passivsummen geben an, welche Kriterien sich über den Programmzeitraum am stärksten verändert haben (Abbildung 6). Positive und negative Werte werden nicht addiert, da sie sich nicht gegenseitig aufheben können, sondern getrennt dargestellt.

6. Entwicklungs- und Wirkungsprofil (Abbildung 7)

Die wichtigsten Informationen aus allen Instrumenten werden darin zusammengefasst. Das Entwicklungsprofil (Trendanalyse) dient der Übersicht, ist aber auch dazu geeignet festzustellen, ob die Gesamtentwicklung robust (gleichmäßig) oder vulnerabel (ungleichmäßig) verläuft. Es wird erkennbar, welche Hauptfaktoren die Entwicklung fördern und welche Rolle Entwicklungsmaßnahmen verschiedener Organisationen sowie externe Einflüsse dabei spielen.

7. MDG-Tabelle

In einer Tabelle werden die MDG mit Unterzielen aufgeführt (erste Spalte) und von der Einflussmatrix und den hierzu gegebenen Erklärungen wird jeweils direkt abgeleitet (zweite Spalte), ob und inwieweit Beiträge zu den MDGs durch die Wirkungen geleistet wurden.

8. Partizipative Entwicklungsplanung

Es werden nun die „Mangelkriterien“ herausgegriffen und hierfür eine Vision entwickelt. Daraufhin werden erste Schritte definiert, wie deren Entwicklung in diese Richtung besser gefördert werden könnte, und es wird festgelegt, wer für diese Schritte als erste/r verantwortlich ist (z. B. wer eine Antragstellung übernimmt). Mit dem Instrument soll erreicht werden, dass der Evaluierungsprozess Ausgangspunkt für weitere Entwicklungsprozesse darstellt.

ABBILDUNG 7: Beispiel eines Entwicklungs- und Wirkungsprofils bei MAPP

	-- Profilinie +/- ++	Kommentare der Diskussionsgruppe / hier Zielgruppen (ZG)	Cross-Checking-Daten	Haupt- einfluss- faktoren
Lebens- standard				
Landwirtsch. Erträge	○ ○ ● ○ ○	Erträge hängen haupts. von den Niederschlägen ab. Durch Maßnahmen schwanken sie weniger.	Im Land sind sinkende Erträge zu beobachten, über die Projektregion gibt es keine Statistiken.	EF/ PORO
Familienein- kommen	○ ○ ○ ○ ●	Das Einkommen steigt aufgrund aller Maßnahmen.	Projektmitarbeiter bestätigen Tendenz in Interviews.	PORO
Gesundheit der Kinder	● ○ ○ ○ ○	Meningitis-Epidemien haben viele Leben gekostet.	Gespräch mit Krankenhauspersonal bestätigt die Todesfälle.	EF/AO
Konsumenten- preise	○ ○ ○ ● ○	Sinkende Preise für uns günstig, da keine volle Selbstversorgung.	Durch Getreidebank können Preisschwankungen abgefedert werden.	EF
Zugang zu				
Feuerholz	○ ○ ○ ○ ●	Durch Eselskarren ist die Reichweite zu Feuerholz gestiegen.	Diese Wirkung ist aus ökologischer Sicht aber sehr negativ.	PORO/ EF
Trinkwasser	○ ○ ○ ○ ●	Durch die Pumpe ist die Nutzung sauberen Brunnenwassers möglich.	Weniger Darmerkrankungen gemäß Krankenhauspersonal.	AO
Märkten	○ ○ ○ ○ ●	Viel besserer Zugang wegen Eselskarren.	Auf den lokalen Märkten wurden viele Menschen aus Sigi getroffen.	PORO
Transport- mitteln	○ ○ ○ ○ ●	Bessere Situation wegen Eselskarren.	PORO subventioniert Eselskarren.	PORO
fruchtbarem Boden	○ ○ ○ ○ ●	Bodenschutzmaßnahmen haben Bodenfruchtbarkeit stark erhöht.	Bodenanalysen auf 10 Äckern bestätigen Aussagen der ZG.	PORO
Wissen				
Kenntnisse	○ ○ ○ ○ ●	Über das PORO und die Kurse haben wir viel gelernt.	Kann aus M&E-Daten nicht direkt nachvollzogen werden.	PORO
Einschu- lungsrate	○ ○ ○ ○ ●	Die Rate beträgt über 90%.	Deckt sich mit den Statistiken.	AO
Rechte und Macht				
Verständigung Geschlechter	○ ● ○ ○ ○	Frauen haben keinen Zugang zu Eselskarren.	Eselskarren erfüllen multiple ungeplante Funktionen.	PORO
LN-Konflikte zw. A./V.	● ○ ○ ○ ○	Die Rechte der V. haben sich verbessert, A. haben weniger Macht.	Verschlechterung des Verhältnisses ist im gesamten Land eingetreten.	EF/ PORO
Verhinderung Landflucht	○ ○ ○ ○ ●	Die Landflucht ist gesunken, weil es wieder eine Perspektive in Sigi gibt.	Und: Keine Einkommensmöglichkeiten für Migranten mehr in Nachbarländern.	EF/ PORO
Legende: EF = Externe Faktoren, AO: Maßnahme einer anderen Organisation, PORO: zu evaluierendes Programm				

b) Akteure

MAPP kann sowohl vom Projektpersonal als auch von einheimischen Fachkräften sowie von externen Gutachtern eingesetzt werden. In der Praxis überwiegt die Fremdevaluierung. Die Zielgruppen nehmen bei MAPP eine zentrale Rolle ein, denn sie sind diejenigen, die die Matrizen mit Hilfe des Evaluators, der eine Moderatorenrolle einnimmt, ausfüllen. Die Cross-Checking-Instrumente werden vom Evaluator bzw. den Evaluationsteams eingesetzt.

c) Datenerhebung

Die Datenerhebung beruht in erster Linie auf den Aussagen der Diskussionsgruppen, die heterogen zusammengesetzt sind, d. h. neben Zielgruppen nehmen auch andere Stakeholder teil (z. B. negativ Betroffene). Konsense werden als solche festgehalten. Kontroversen bzw. Dissense werden ebenfalls identifiziert, diskutiert und als solche festgehalten, soweit sie systematisch auftreten und für soziale Sub-Gruppen gültig sind (z. B. für besonders arme Bevölkerungsgruppen oder Frauen). Es können gesonderte Punktwerte für Sub-Gruppen vergeben werden, indem z. B. für kontroverse Beurteilungen verschiedene Farben für die Punkte in den Matrizen verwendet werden.

Sollen die Wirkungen größerer Programme erhoben werden, dann müssen mehrere MAPP-Durchgänge mit mehreren Diskussionsgruppen durchgeführt werden, um die Bandbreite der Wahrnehmungen aufzufangen. Die Definition der Stichprobe erfolgt in der Regel qualitativ indem eine bewusste Auswahl entlang von Kriterien getroffen wird. Allerdings steht prinzipiell nichts einer Zufallsauswahl, d. h. repräsentativen Stichprobe, entgegen.

d) Datenauswertung

Der erste Schritt der Datenauswertung erfolgt nach einem festen Muster (siehe Instrument 6: Entwicklungs- und Wirkungsprofil unter 4a) gemeinsam mit den Diskussionsgruppen. Im Evaluationsreport werden die Ergebnisse der unterschiedlichen MAPP-Sitzungen dann reflektiert, u. U. aggregiert und zusammengefasst. Treten unterschiedliche Wahrnehmungen oder Widersprüche zwischen Selbst- und Fremdwahrnehmung auf (Wahrnehmung der Zielgruppen und Wahrnehmung des Evaluationsteams) oder Differenzen mit den Cross-Checking-Daten, werden diese am Schluss mit den Diskussionsgruppen nochmals thematisiert und dargelegt. Echte Widersprüche kommen insgesamt sehr viel seltener vor, als man erwarten dürfte. Manche Differenzen lösen sich in der Debatte auf. Oft stellt sich heraus, dass sie nur unterschiedliche Aspekte des gleichen Sachverhalts beschreiben.

5. Nutzung der Ergebnisse

Die Ergebnisse sind leicht verständlich und geben transparent die Sichtweise der Stakeholder wieder. Die daraus zu ziehenden Schlussfolgerungen und evtl. neue Maßnahmen können mit Hilfe der partizipativen Entwicklungsplanung (Instrument 8) entwickelt werden und als Vorschläge direkt in das evaluierte und/oder in das neu aufzulegende Programme eingebracht werden. Gegebenenfalls identifizierte negative Wirkungen können Ausgangsbasis für die Entwicklung von Kompensationsleistungen sein oder Anstoß zu Umsteuerungen darstellen.

Je nachdem, wie groß die Programme und die Stichproben sind, die durch das Evaluierungsteam befragt wurden, können die Ergebnisse von MAPP aggregiert werden, so dass auch flächendeckende Maßnahmen evaluiert werden können.

6. Anforderungen

Ein MAPP sollte organisatorisch durch die Vor-Ort-Mitarbeiter vorbereitet werden. Die reine Erhebungs- und Auswertungszeit vor Ort dauert rund 2-3 Tage, in Abhängigkeit von der Anzahl der eingesetzten Instrumente sowie der Anzahl der Kriterien und Maßnahmen, die bewertet werden. Werden mehrere MAPP-Durchgänge durchgeführt (größere Stichprobe aufgrund flächendeckender Programme) und/oder Untergruppen gebildet (z. B. Frauen- und Männergruppen), verlängert sich die Zeit entsprechend.

MAPP-Analysen können durch geschulte einheimische Fachkräfte durchgeführt werden, wenn diese Moderationserfahrung und/oder Erfahrungen in partizipativen Methoden aufweisen.

Voraussetzung für die erfolgreiche Durchführung ist eine gewisse Diskussionskultur im entsprechenden Partnerland, so dass sowohl echte Konsense als auch kontroverse Wahrnehmungen in den Gruppendiskussionen erkannt werden können.

7. Besondere Eignung und Grenzen des Verfahrens

Mit MAPP können alle diejenigen Maßnahmentypen untersucht werden, für die Zielgruppen klar identifiziert und deren Wirkungen von ihnen wahrgenommen werden können. Geeignet ist MAPP insbesondere für die Evaluierung multidimensionaler Zielkonzepte wie Armutsminderung, Demokratisierung, Dezentralisierung etc., deren Beschreibung am Besten anhand mehrerer Kriterien (multidimensional) vorgenommen wird. Die Überbrückung der „Zuordnungslücke“ wird bei MAPP explizit vorgenommen. Hierbei be-

stehen auch Möglichkeiten der Gegenüberstellung verschiedener Sichten und Perspektiven. Die systematische Einbeziehung der Nutzer erleichtert eine direkte Nutzung der Ergebnisse und eine Identifizierung des Bedarfs für weitergehende Maßnahmen.

Das Instrument fokussiert auf die Wirkungsebene (*outcome* und *impact*). Ist es allerdings Ziel, auch die qualitative Ausgestaltung von Maßnahmen (Input, Output) zu erfassen, sollte MAPP mit anderen Methoden kombiniert werden. Die Qualität und Validität der Ergebnisse von MAPP hängen von der Moderationsfähigkeit des Evaluationsteams, u. U. der Qualität des Übersetzers / der Übersetzerin und der Zusammensetzung der Diskussionsgruppe ab. Auf ausreichend Zeit beim Diskutieren und auf eine heterogene Zusammensetzung der Diskussionsgruppen sollte geachtet werden.

Literaturhinweise

- DEZA/Seco (2008): Wirkungsbericht der Schweizer Entwicklungszusammenarbeit im Wassersektor. Bern. [Beispiel einer Sektorevaluierung mit Hilfe von MAPP]
- Neubert, S. (1998): SWAP (MAPP) – ein neues System zur Wirkungsanalyse armutsorientierter Projekte in der Entwicklungszusammenarbeit. Berlin: DIE. URL: <http://www.cefe.net/forum/neubert.pdf> (11.1.09)
- Neubert, S. (2004a): Akteurszentrierte Wirkungsanalyse in der Entwicklungszusammenarbeit – ein Leitfaden. In: ded-Brief 3/2004. Bonn: DED. [Praktischer Leitfaden zur Durchführung]
Link: http://www.ded.de/cipp/ded/lib/all/lob/return_download,ticket.g_u_e_s_t/bid_534/no_mime_type,0/~DEDBrief_04_3_Qualitaet_und_Wirkung.pdf (11.1.09)
- Neubert, S. (2004b): Akteurszentrierte Wirkungsanalyse – konzeptionelle Überlegungen und *Scaling up* für die Verwaltungszusammenarbeit. In: K. Simon (Hrsg.): Verwaltungen, Experten und Bürger im Reformprozess – Wirkungen und Evaluierung von Verwaltungszusammenarbeit mit Entwicklungsländern. Baden-Baden: Nomos. [Konzeptionelles]
- Neubert, S. (2004c): Wirkungsanalysen der entwicklungspolitischen Zusammenarbeit sind machbar. Analysen und Stellungnahmen 4/2004. Bonn: DIE. [Einordnung in die Debatte]

Most Significant Change (MSC)

STEFAN DOFEL

1. Allgemeine Informationen

Most Significant Change (MSC) ist eine qualitative, interpretierende Analysetechnik zur Feststellung des bedeutsamsten herbeigeführten Wandels und/oder unterschiedlicher Dimensionen des Wandels im Kontext einer Entwicklungsmaßnahme (Projekt, Programm oder Ansatz / SWap). MSC beruht dabei auf einem partizipativen Ansatz, der den Lerneffekt über Wirkungsweisen in den Vordergrund rückt.

Technik und methodisches Vorgehen wurden 1994 im Rahmen einer Feldstudie über organisationelles Lernen in NRO von Rick Davies in Bangladesch entwickelt. Motiviert wurde die Entwicklung durch die Herausforderung, Monitoring und Evaluation eines komplexen, vielschichtigen Programms zur ländlichen Entwicklung zu leisten, das in über 700 dörflichen Gemeinden parallel implementiert wurde. Seit 1994 wurde MSC von DFID, SIDA und vor allem einer ganzen Reihe von internationalen NRO (Oxfam, ADRA, CARE, CBM u. a.) zur Wirkungserfassung und zum organisationellen Lernen sowohl in Entwicklungs- als auch Industrie- und Schwellenländern eingesetzt.

2. Anwendungsbereiche

MSC ist insbesondere geeignet für vielschichtige und komplexe Maßnahmen sowie für Maßnahmen, die auf Armutsbekämpfung, sozialen Wandel und/oder die Bereitstellung sozialer Dienstleistungen abzielen. Anwendungserfahrungen liegen vor allem in den Sektoren ländliche Entwicklung, Bildung und Gesundheit vor. Das indikatorenfreie Verfahren kann während der Maßnahmendurchführung (Monitoring) ebenso zum Einsatz kommen wie im Rahmen von Evaluierungen zum Abschluss einer Maßnahme oder ex-post. Mit Hilfe von MSC werden nicht nur Zustände erfasst, sondern ein detailliertes Verständnis von Wirkungszusammenhängen und relevanten Faktoren (Prozessen) erzeugt. MSC kann zur Erfassung von Wandel und

Wirkungen auf allen gesellschaftlichen Ebenen (Mikro-, Meso- und Makroebene) dienen.

3. Analyseansatz

Bei MSC werden konkrete Fallbeispiele / Ereignisse auf Zielgruppen- und Organisations-/Partnerebene über herbeigeführten Wandel im Kontext einer Entwicklungsmaßnahme mittels offener Interviews (Geschichtenerzählungen) gesammelt und analysiert. Auf diese Weise können direkte und indirekte, beabsichtigte und unbeabsichtigte, positive wie negative Wirkungen erfasst werden.

Von zentraler Bedeutung ist an dieser Stelle die Vorabauswahl von breit gefassten Dimensionen der Veränderung, die ermittelt werden sollen. Dadurch wird bei der jeweiligen Konzipierung des Monitorings oder der Evaluation festgelegt, welche Blickrichtung bei Anwendung der Technik eingenommen wird. Die Zuordnung von Wirkungen wird dann entweder mittels gezielter Nachfrage durch den/die Interviewer/-in geleistet, oder im Rahmen von Validierungen durch tiefer gehende Prüfung einzelner Fallbeispiele / Ereignisse (Geschichten) und deren Ursachen („Brutto- vs. Nettowirkungen“).

In theoretischer Hinsicht ist MSC dem Ansatz des Appreciative Inquiry (AI) verwandt, geht jedoch über diesen hinaus durch Berücksichtigung auch negativer und unerwarteter Wirkungen. Im Gegensatz zu AI ist MSC aber weniger für die Planungsphase einer Intervention, sondern eher für Monitoring und/oder Evaluation geeignet.

4. Durchführung

a) Verfahrensablauf

MSC umfasst zehn Schritte:

1. Start und Wecken von Interesse

Im ersten Schritt werden betroffene Interessengruppen / zukünftige Anwender in die MSC-Technik eingeführt und ihr Interesse für ein engagiertes Mitwirken geweckt.

2. Definition von Veränderungsbereichen

Weit gefasste Wirkungsbereiche wie „Veränderungen in Lebensbedingungen“ werden von ausgewählten Interessenvertretern festgelegt. Eine weiter-

gehende Spezifizierung wird im Monitoringprozess (bzw. Evaluierungsprozess) später von den tatsächlichen Anwendern vollzogen.

3. Festlegung des Beobachtungszeitraums

Es wird vereinbart, über welche Zeiträume und wie häufig Veränderungen der unter 2. festgelegten Wirkungsbereiche beobachtet werden sollen.

4. Sammlung von Geschichten zu wesentlichen Veränderungen

Bei den Personengruppen, die eng mit dem definierten Wirkungsbereich in Verbindung stehen, werden mit Hilfe einfacher Fragetechniken Geschichten zu beobachteten wesentlichen Veränderungen gesammelt. Zudem wird erfragt, warum die genannte beschriebene Veränderung als besonders bedeutend eingeschätzt wird.

5. Auswahl der wichtigsten und signifikanten Geschichten

Die Geschichten werden dann von den einzelnen Hierarchiestufen der betroffenen Förderorganisation analysiert und gefiltert. Jede einzelne Ebene erhält Geschichten von der unter ihr befindlichen Hierarchiestufe. Sie überprüft die erhaltenen Geschichten, wählt die für sie wichtigsten aus und leitet sie an die ihr vorgesetzte Hierarchieebene weiter. So wird die Anzahl der Geschichten in einem transparenten Prozess reduziert.

6. Rückspiegeln der Ergebnisse des Auswahlprozesses

Die bei der Auswahl der Geschichten zum Einsatz kommenden Kriterien werden allen am Auswahlprozess Beteiligten mitgeteilt (begründete Auswahl!). Die Organisation macht sich so ihren eigenen Aufmerksamkeitsfokus und ihre Selektionskriterien bewusst und erleichtert eine bewusste Steuerung der jeweiligen Entwicklungsmaßnahme. Ein zusammenfassendes Dokument wird erstellt, das neben den ausgewählten Geschichten auch die Begründungen für deren Auswahl umfasst. Vertreter von Geberorganisationen werden dann gebeten, die Geschichten zu bewerten und jene auszuwählen, die am besten die von ihnen besonders erwünschten Veränderungen beschreiben. Zusätzlich legen sie offen, warum sie die jeweiligen Geschichten ausgewählt haben. Geschichten, die negative Veränderungen beschreiben, werden analysiert und ggf. als Begründung für eine Anpassung der Intervention genommen.

7. Überprüfung der Geschichten (Validierung)

Ausgewählte Geschichten können vor Ort überprüft werden. Dabei wird eine doppelte Zielrichtung verfolgt: zum einen wird der Wahrheitsgehalt der

Geschichten überprüft; zum anderen werden zusätzliche Informationen zu den beschriebenen Veränderungen gesammelt.

8. *Quantifizierung*

Die Veränderung wird quantitativ erfasst und, wenn möglich, mit einer vergleichbaren Situation an einem anderen Ort verglichen. Alternativ können thematische Codes für einzelne Veränderungen eingeführt werden, nach welchen sämtliche Geschichten überprüft werden; die Zählung der Nennungen der einzelnen Veränderungen erlaubt die Quantifizierung der qualitativen Daten.

9. *Sekundäranalyse und Meta-Monitoring*

Das Monitoring bzw. die Evaluation wird einer kritischen Analyse unterzogen. Dabei wird u. a. die qualitative Beteiligung von Interessengruppen oder die Häufigkeit der Nennung bestimmter Veränderungen untersucht.

10. *Veränderung / Anpassung des Systems*

Auf Grundlage der gemachten Erfahrungen und Analysen wird das Monitoringsystem bzw. das Evaluationsverfahren auf Grundlage der MSC-Technik angepasst.

Zur Einführung der MSC-Technik als Bestandteil des regulären Monitoringsystems einer Entwicklungsmaßnahme ist eine rigide Implementierung anhand vorab festgelegter zeitlicher Abstände und Modalitäten der Sammlung, Analyse und Weiterreichung der ausgewählten Fallbeispiele / Ereignisse (Geschichten) an die nächst höhere Hierarchieebene erforderlich. Ebenso ist ggf. eine Schulung der Mitarbeiter/-innen auf Gemeindeebene notwendig zur kundigen Sammlung und Dokumentation der Fallbeispiele und Ereignisse (Befragungstechniken, ethische Implikationen, etc.). Schablonen für die gleichförmige Dokumentation der Fallbeispiele (Datenerfassungsprotokolle / *data capture protocol*) sowie für die Begründung der Auswahl zur Weitergabe an die nächst höhere Hierarchieebene sind vorab zu erstellen.

b) Akteure

Für die Durchführung von MSC, speziell die Arbeitstreffen bzw. die partizipativen Workshops zur Auswahl von Fallbeispielen / Ereignissen (Geschichten), ist eine externe Moderation wünschenswert, aber nicht zwingend erforderlich. Zielgruppen und Mitarbeiter/-innen aller relevanten Hierarchieebenen sind aktiv in den gesamten Prozess – Datensammlung, Diskussion und

Auswertung – einbezogen. So wird auch deren Kompetenz in Wirkungsidentifizierung und -analyse gestärkt.

c) Datenerhebung

Datenquellen für MSC sind Personen (Zielgruppenvertreter/-innen, Mitarbeiter/-innen), die nicht zufällig ausgewählt, sondern aussagefähig zu Veränderungen in bestimmten Wirkungsbereichen (zweckmäßige Befragung) bzw. repräsentativ für die Zielgruppe sind (Repräsentativbefragung). Zentrales Erhebungsinstrument ist das narrative (v. a. Gruppen-)Interview. Abhängig vom Ziel des Einsatzes der MSC-Technik können entweder unterschiedliche Perspektiven / Wahrnehmungen im Prozess der Analyse und Filterung von Fallbeispielen / Geschichten berücksichtigt werden, oder ein einziges Fallbeispiel / Geschichte am Ende des Auswahlprozesses identifiziert werden, dem repräsentative oder besonders signifikante und aussagestarke Eigenschaften zugeordnet werden. In jedem Fall ist die partizipative Analyse und Reflektion über Fallbeispiele / Geschichten im Rahmen des Auswahlprozesses zentraler Baustein der MSC-Technik. Form der erhobenen Daten ist ausschließlich Text, der mittels vorab erstellter Datenerfassungsprotokolle gleichförmig dokumentiert wird.

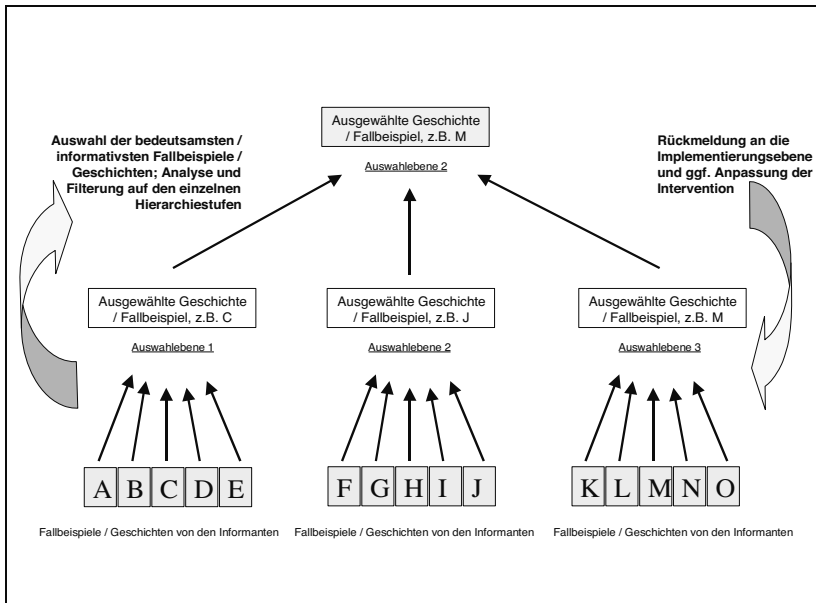
d) Datenauswertung

Die Auswertung (Analyse, Filterung) der Fallbeispiele / Geschichten erfolgt anhand festgelegter Kriterien in den Schritten 5 und 6, die Validierung der Ergebnisse – durch Vor-Ort-Überprüfung – explizit in den Schritten 7 und 8 des Verfahrens (siehe oben). Darüber hinaus wird die MSC-Anwendung (im Rahmen von Monitoring oder Evaluierung) selbst kritisch reflektiert.

5. Nutzung der Ergebnisse

Die Ergebnisse von MSC sind für alle Beteiligte in hohem Maße transparent. Auch ist eine Aggregation der Daten (bzw. Geschichten) im Verfahren angelegt und Quantifizierung – mit Einschränkungen – möglich. Die Nutzung der Ergebnisse für Steuerung und Lernen – bezogen auf die Entwicklungsmaßnahme wie auch auf das Monitoring- und Evaluierungssystem – ist integraler Bestandteil des Verfahrens.

ABBILDUNG 8: Die MSC-Technik im Überblick



Quelle: In Anlehnung an Davies, R. and Dart, J. J. (2005)

6. Anforderungen

Zur Einführung von MSC als Bestandteil des Monitoringsystems ist ein erheblicher zeitlicher und personeller Aufwand erforderlich (Sammlung und Dokumentation konkreter Fallbeispiele / Ereignisse sowie regelmäßige, voneinander unabhängige Arbeitstreffen der Mitarbeiter/-innen aller relevanten Hierarchieebenen). Regelmäßige Arbeitstreffen zur Auswahl der bedeutsamsten und informativsten Fallbeispiele / Ereignisse (Geschichten) finden idealerweise vierteljährlich statt, können allerdings auch monatlich oder halbjährlich durchgeführt werden.

Zur einmaligen Anwendung des Verfahrens im Rahmen einer (internen, externen oder gemischten) Wirkungsevaluierung ist die Sammlung und Dokumentation einer ausreichenden Anzahl von Fallbeispielen / Ereignissen (Geschichten) im Vorfeld der Evaluierung erforderlich, die dann im Rahmen eines eintägigen, partizipativen Workshops diskutiert und analysiert werden können. Vor- und Nachbereitung erfordern insgesamt zusätzlich ca. 3 Tage.

Folgende Voraussetzungen werden für den erfolgreichen Einsatz von MSC als hilfreich angesehen:

- Eine Organisationskultur, die es ermöglicht, Erfolge genauso wie Fehlschläge offen zu besprechen;
- einflussreiche Förderer in Führungspositionen mit guten Vermittlungsfähigkeiten;
- Bereitschaft, neue Ansätze auszuprobieren;
- genügend Zeit, mehrere Zyklen des Ansatzes durchzuspielen;
- Möglichkeiten, den beteiligten Interessensgruppen ein regelmäßiges Feedback zukommen zu lassen;
- Unterstützung und Veränderungsbereitschaft der Führung der Organisation;
- Fokussierung auf soziale Veränderungen;
- partizipative Grundausrichtung der Organisation.

7. Besondere Eignung und Grenzen des Verfahrens

MSC erleichtert die Identifizierung unerwarteter Wirkungen und ist besonders dienlich bei Entwicklungsmaßnahmen, deren Wirkungen schwierig zu quantifizieren sind. Dabei schafft MSC einen reichhaltigen Überblick und reduziert eine komplexe Wirklichkeit nicht auf eine einzige Kennziffer. Das Verfahren hilft, die Werte einer Organisation transparent zu machen und sie systematisch zu reflektieren bzw. kritisch zu analysieren. Die Anwendung von MSC erfordert keine besonderen fachlichen Voraussetzungen, auch erscheint die Vermittlung über kulturelle Grenzen hinweg unproblematisch; allerdings verlangt die Datenauswertung bzw. -analyse einige Kompetenz. Eine besondere Stärke von MSC liegt darin, dass es auch für Monitoring- und Evaluierungsaktivitäten eingesetzt werden kann, bei denen nicht bereits im Vorfeld die zu erreichenden Veränderungen definiert sind.

Grenzen des Verfahrens liegen in der Notwendigkeit einer kontinuierlichen Mitwirkung einer großen Zahl von Mitarbeiter/-innen. Auch ist für Entwicklungsmaßnahmen, die über keine Strukturen (Mitarbeiter/-innen) auf Zielgruppenebene verfügen, die Datensammlung sowie die Validierung einzelner Geschichten problematisch.

Literaturhinweise

- Dart, J. J.; Drysdale, G.; Cole, D. und Saddington, M. (2000): The Most Significant Change approach for monitoring an Australian Extension project. In: PLA Notes, 2000, Vol. 38, S. 47-53, London: International Institute for Environment and Development. <http://www.clearhorizon.com.au/page.php?pid=10700&wpid=7> (24.11.2008)
- Dart, J. J. und Davies, R. (2003): A Dialogical, Story-Based Evaluation Tool: The Most Significant Change Technique. In: American Journal of Evaluation, 2003/24, S. 137-155.
- Davies, R. und Dart, J. J. (2005): The 'Most Significant Change' (MSC) Technique. A Guide to Its Use. Cambridge: Self-published. <http://www.mande.co.uk/docs/MSCGuide.htm> (11.11.08)
- Davies, R. (1996): An Evolutionary Approach to Facilitating Organisational Learning: An Experiment by the Christian Commission for Development in Bangladesh. Centre for Development Studies. In: Mosse, David; Farrington, John und Rew, Alan (Hrsg.) (1998): Development as Process: Concepts and Methods for Working with Complexity. Overseas Development Institute (ODI) Policy Studies. London: Routledge. (Also published in: Impact Assessment and Project Appraisal, September 1998 Vol. 16, No. 3. S. 243-250). <http://www.mande.co.uk/docs/ccdb.htm> (11.11.08)
- Le Cornu, Rosie; Peters, Judy; Foster, Margot; Barratt, Robyn und Mellowship, Diane (2003): Exploring Perceptions of 'Significant Change' in Reforming Schools. Paper presented to: NZARE/AARE Joint Conference, Auckland, November 30th December 3rd, 2003. (Subsequently published in "Perceptions of 'Significant Change' in School Cultures in South Australia" by Rosie Le Cornu, Judy Peters, Margot Foster, Robyn Barratt and Jacqueline Stratfold, International Journal of Knowledge, Culture and Change Management, Volume 6, Issue 5, S. 161-170).
- Lunch, Chris (2007): The Most Significant Change: using participatory video for monitoring and evaluation. In: Participatory Learning and Action, Volume 56, Number 1, June 2007, S. 28-32(5). <http://www.insightshare.org/pdfs/PLA%20PV%20ARTICLE%2007.pdf>
- Sigsgaard, P. (2004): Doing away with Predetermined Indicators: Monitoring Using the Most Significant Changes Approach. In: Earle, L. (Hrsg.) Creativity and Constraint. Grassroots Monitoring and Evaluation and the International Aid Arena. NGO Management & Policy Series 18. Oxford: INTRAC. http://www.petersigsgaard.dk/PDFfiler/doing_away_with_%20predetermined_indicators.pdf (24.11.08)
- Willets, Juliet und Crawford, Paul (2007): The most significant lessons about the Most Significant Change technique. In: Development in Practice, June 2007, Vol. 17, Issue 3, S. 367-379.

NGO-IDEAs Impact Toolbox

KIRSTEN VORWERK / EBERHARD GOHL

1. Allgemeine Informationen

Die NGO-IDEAs Impact Toolbox wurde im Rahmen eines Projektes zur Ermittlung der Wirkungen von Spar- und Kreditprogrammen südindischer NRO entwickelt. Am Projekt NGO Impact on Development, Empowerment and Actions (NGO-IDEAs) waren 14 deutsche NRO und 32 indische Partnerorganisationen beteiligt (www.ngo-ideas.net). Folgeprojekte werden von Oktober 2008 bis Dezember 2011 in verschiedenen Sektoren in Südasien (Indien, Bangladesch), Philippinen und Ostafrika (Kenia, Tansania, Uganda) durchgeführt.

Die NGO-IDEAs Impact Toolbox beabsichtigt in erster Linie, (Selbsthilfe)-Gruppen (SHG, CBO) und NRO bei der wirkungsorientierten Steuerung zu unterstützen, indem sie Veränderungen in der Situation der einzelnen Personen, Haushalte und Gruppen sichtbar (beschreibbar und messbar) macht und Möglichkeiten zu Aggregation und Disaggregation, Analyse und Reflexion aufzeigt. Zugleich aber ist das Verfahren mit den ergänzenden Erhebungen und Analysen und seinem Berichtsformular darauf angelegt, auch die Rechenschaftslegung der NRO zu verbessern.

2. Anwendungsbereiche

Die NGO-IDEAs Impact Toolbox wurde zunächst entwickelt für Projekte mit Mikrofinanzkomponente (Sparen und Kredit), wird allerdings ausgedehnt auf weitere Sektoren wie Gesundheit (Basisgesundheitsdienste) und Bildung (*empowerment*). Als Querschnittsthemen wurden Gender, Überwindung von Armut sowie Behinderung ausgewählt.

Die Toolbox ist geeignet für jegliche Förderung von Selbsthilfe (z. B. Beratung und Training) durch einen Gruppenansatz. Sie dient der Ermittlung von Veränderungen auf der Mikroebene, bei Einzelpersonen, Haushalten und Gruppen. Im Wesentlichen handelt es sich um ein periodisches Monitoring-Verfahren für laufende Projekte. Durch die damit verbundene Feststellung der Ziele der jeweiligen Gruppe kann es auch als Planungsins-

trument gesehen werden. Der Schwerpunkt des Ansatzes liegt auf der Erfassung von Veränderungen von Zuständen.

3. Analyseansatz

Neben Veränderungen bei Einzelpersonen oder Familien in den von den Zielgruppen selbst definierten Bereichen werden Veränderungen bei Gruppen, basierend auf den Zielen, die sich die (Selbsthilfe-) Gruppe gesetzt hat, erfasst. Darüber hinaus werden unvorhergesehene positive oder negative Veränderungen untersucht sowie Auswirkungen auf das Geschlechterverhältnis (*Gender Disaggregation*) und Veränderungen bei unterschiedlichen Wohlstands-/Armuts-Gruppen (*Poverty Disaggregation*). Betrachtet werden insbesondere kurz- und mittelfristige Wirkungen.

Der Nachweis erfolgt über festgelegte qualitative und quantitative Indikatoren über einen Vorher- (zu Beginn des Projektes) / Nachher- (zum Erhebungszeitpunkt) Vergleich. Dem Zuordnungsproblem wird dadurch begegnet, dass die NRO

- die projektinternen und -externen Faktoren systematisch erfasst, die zu den beobachteten Veränderungen beigetragen haben, sowie eine Einschätzung von deren Bedeutung für die beobachteten Veränderungen vornimmt (je nach Befähigung erfolgt dies auch durch die Gruppenmitglieder);
- eine Untersuchung von Vergleichsgruppen (z. B. Vergleich mit staatlich geförderten Selbsthilfegruppen) und/oder von Selbsthilfegruppen mit unterschiedlich langer Förderung unternimmt.

4. Durchführung

a) Verfahrensablauf

Folgende Instrumente sind in der Toolbox aufgeführt:

- PWR – Participatory Wealth Ranking gruppiert die Haushalte nach Wohlstands-/Armuts-Merkmalen:
 - Es dient der armutsorientierten Zielgruppenauswahl und der Analyse von Veränderungen bei unterschiedlichen sozialen Gruppen (sehr Arme / Arme / weniger Arme).
- SAGE – Situational Analysis and Goal Establishment erfasst Veränderungen bei Einzelpersonen und Familien:

- Dieses in NGO-IDEAs entwickelte Instrument baut auf den Zielen auf, die sich die jeweiligen SHG geben, und die überwiegend die Vision der Gruppenmitglieder von ihren künftigen Lebensumständen beschreiben. Aus den Zielen werden unmittelbar Indikatoren abgeleitet. Mit diesen Indikatoren können die Gruppenmitglieder die von ihnen angestrebten Veränderungen (z. B. bezogen auf Arbeit, Einkommen, Ernährung, Ausbildung, familiäre Gewalt, Engagement für die Gemeinschaft) selbst beobachten und bewerten.
- PAG – Performance Appraisal of the Groups erfasst Veränderungen von Gruppen:
 - Es basiert auf dem Ratingverfahren, das die Banken bei der Kreditvergabe an SHG anwenden. Das Verfahren ist schwerpunktmäßig auf ökonomische Veränderungen ausgerichtet. Mit NGO-IDEAs wurde dieses Verfahren um nicht-ökonomische Merkmale der Gruppe erweitert. Ein durch die SHG selbst, ggf. mit der NRO, durchgeführtes Bewertungsverfahren dient der Wirkungserfassung auf Gruppenebene.
- PAngo – Performance Assessment of NGOs dient der vertieften Analyse der Wirkungen anhand mehrerer Elemente:
 - Zusammenstellung der Daten der quantitativen Indikatoren aus SAGE, PAG und anderen Quellen;
 - Zusammenstellung der Daten der deskriptiven Indikatoren aus SAGE, PAG und anderen Quellen;
 - Vergleich der Veränderungen mit Vergleichsgruppen;
 - Analyse der Wirkungszusammenhänge;
 - Gender-spezifische Disaggregation der Wirkungen;
 - Armuts-spezifische Disaggregation der Wirkungen: Sofern die Personen und Haushalte aufgrund des PWR den Wohlstands-/Armutsmerkmalen zugeordnet werden können, kann z. B. sichergestellt werden, dass die Projektwirkungen auch bei Menschen in extremer Armut ankommen.

b) Akteure

Die Methoden SAGE und PAG sind so konzipiert, dass sie von den Gruppen-Mitgliedern selbst angewandt werden können. Zu Beginn sind jedoch die Anleitung und Moderation durch das begleitende NRO-Personal notwendig. PAngo ist als Monitoringsystem der NRO angelegt, welches auf SAGE und PAG aufbaut. Die gesamte Impact Toolbox dient in erster Linie der Erfahrung der eigenen Wirksamkeit von Zielgruppen und NRO.

ABBILDUNG 9: Interne Kontrollmechanismen zur Validierung von Ergebnissen bei NGO-IDEAs

Interne Kontrollmechanismen				
				
Selbstbeurteilung durch die einzelnen Gruppen-Mitglieder	Selbstbeurteilung und Validierung durch die SHG	Beurteilung der SHG und Validierung durch die <i>Apex Bodies</i>	Beurteilung und Validierung durch die NRO	Externe Beurteilung und Validierung
Die Grundprinzipien des partizipativen Ansatzes werden als Teil der internen Kontrollmechanismen angewandt, auf der Grundlage von transparenten Methoden, die auch der Rechenschaftspflicht Genüge tun.				
Jedes Instrument hat eine „Triangulation“ (= eine Betrachtung aus anderen Perspektiven) eingebaut.				

c) Datenerhebung

In erster Linie erfolgt die Datenerhebung durch die Zielgruppen selbst (mit SAGE und PAG), darüber hinaus bei PANGo auch über andere Quellen und Sekundärmaterial. Eine Stichprobenauswahl ist in der Methodik nicht angelegt, da sie in erster Linie für das eigene Wirkungsmonitoring der Gruppen entwickelt wurde. Sofern die Impact Toolbox nur zur Erhebung eingesetzt werden soll, sollten die Auswahlkriterien sich auf die Gruppen beziehen (z. B. nach Regionen, sozialen Merkmalen, Dauer des Bestehens, Umfang der ihnen angebotenen Projektleistungen), nicht auf Personen oder Haushalte, weil der Ansatz die gleichartige Mitwirkung aller Gruppenmitglieder voraussetzt.

Instrumente der Datenerhebung sind Gruppendiskussionen, Einzelbefragungen und Befragungen von Haushalten (standardisiert). Grundlegende Kriterien für die einzelnen Instrumente sind:

- Einfach anzuwenden für SHG und NRO;
- effizient bezogen auf Kosten- und Zeitaufwand;
- partizipativ:
 - Auf dem Wissen der beteiligten Menschen aufbauend;
 - eine Kultur des Lernens und des Erkundens der eigenen Wirksamkeit fördernd;
 - von den Zielgruppen angeeignet werdend (*ownership*);
 - demokratisch funktionierend.

Durch das Analyseinstrument PANGo werden unterschiedliche Perspektiven erfasst, außerdem kann eine Differenzierung nach Geschlecht und/oder so-

zialem Status vorgenommen werden. Darüber hinaus sind eine Reihe von internen Kontrollmechanismen (vgl. Abbildung 9) vorgesehen.

d) Datenauswertung

Die Impact Toolbox sieht eine einfache statistische Auswertung mit Hilfe von Microsoft Excel® vor, eine spezielle Software wird derzeit entwickelt. Bei der Auswertung der personen- bzw. haushaltsbezogenen Daten (SAGE) können durch die Kategorisierung von Personen nach Wohlstands-/Armutsmerkmalen (PWR) auch Aussagen bezüglich der Armutsminderung getroffen werden. Das PANgo-Format dient als Raster für die Auswertung.

Die Analyse der Daten umfasst eine gemeinsame Reflektion der Ergebnisse durch Zielgruppen und NRO. Die Befragung der Gruppen dient (auch) der Validierung der Einzelerhebungen. Triangulation von Methoden und Informationen ist im Rahmen des Verfahrens teilweise angelegt und wird z. T. durch die beteiligten NRO bestimmt. Darüber hinaus gibt es interne Kontrollmechanismen durch die Förderativen Strukturen der SHG sowie anderer Beteiligter (vgl. Abbildung 9).

5. Nutzung der Ergebnisse

Das Verfahren beinhaltet ein Berichtsformular, in dem Ergebnisse, Methoden, Datenquellen, Beteiligte, Feedbackprozess und gezogene Schlussfolgerungen transparent dargelegt werden. Ergebnisse und Schlussfolgerungen werden von den Beteiligten (NRO, Zielgruppen) gemeinsam reflektiert. Die Ergebnisse sind für alle Beteiligten gut nachvollziehbar und handlungsorientiert.

Die Daten sind überwiegend quantifiziert (Angaben in Prozentzahlen), zur Aggregation werden Indikatoren gebündelt. Bei den Erhebungen auf Gruppenebene (PAG) erfolgt eine Vergabe von Punkten (*rating*). Im PANgo Format werden die erhobenen Daten nochmals zusammengefasst und ausgewertet. Aggregation ist aber nur dann möglich, wenn die Gruppen die gleichen Indikatoren verwenden.

Die Toolbox enthält eine Anleitung zur Auswertung der Ergebnisse und Empfehlungen zur Reflexion. Dazu sind im PANgo-Format verschiedene Fragen aufgeführt. Die Auswertung der Daten erfolgt teilweise gemeinsam mit den Zielgruppen, teilweise durch die NRO. Der gemeinsame Reflektionsprozess umfasst auch Formulierung von Schlussfolgerungen und konkreten Handlungsmaßnahmen.

6. Anforderungen

Die Anwendung der Instrumente SAGE und PAG (Zielgruppen / NRO) dauert je nach Intensität der Analyse drei Stunden bis einen Tag. Die konsequente Anwendung aller Tools erfordert zusätzlichen Aufwand bei den NRO für die Ausfüllung des PANgo-Formats (offene Fragen, Heranziehen anderer Datenquellen etc.), Armutsklassifizierung (einmalig am Anfang) sowie Vergleichsgruppenbefragungen und Differenzierung nach Gender. Die Erhebungen sollten im jährlichen Rhythmus durchgeführt werden.

Insgesamt ist das Verfahren sehr kostengünstig, da es einfach in den normalen Arbeitsablauf der NRO integriert werden kann. NRO brauchen allerdings die notwendigen technischen Voraussetzungen zur Datenanalyse (z. B. Computersoftware und -kenntnisse).

Die Methoden SAGE und PAG können leicht erlernt werden (nach ca. 3 Anwendungen von der Zielgruppe durchzuführen), zu Beginn ist jedoch die Anleitung und Moderation durch die NRO notwendig. PANgo ist als Monitoringinstrument der NRO angelegt, welches auf SAGE und PAG aufbaut.

Für die Anwendung der Toolbox (insbesondere von PANgo) sind Training und Beratung von NRO-Personal erforderlich, insbesondere im Bereich Datenauswertung und Aggregation.

7. Besondere Eignung und Grenzen des Verfahrens

Die NGO-IDEAs Impact Toolbox bietet teilweise einfach zu handhabende Instrumente, welche es erlauben, dass NRO und Zielgruppen *ownerships* darüber entwickeln. Gemeinsame Reflektionsprozesse unterstützen die Wirkungsanalyse und Handlungsorientierung. Die Indikatoren, welche die Kernthemen und selbst gesetzten Ziele der Gruppen betreffen, erlauben auch eine Quantifizierung der Wirkungen.

Derzeitige Herausforderungen bestehen darin, dass die Datenanalyse und -aggregation noch Schwierigkeiten bereiten, dass die Indikatoren teilweise nicht SMART sind (spezifisch, messbar, erreichbar/*attainable*, relevant und tzeitgebunden/*timebound*), vorhandene externe Datenquellen noch nicht ausreichend genutzt werden und Vergleichsgruppenverfahren bislang nicht ausreichend getestet wurden.

Literaturhinweis

Karl Kübel Foundation for Child and Family (2007): The Impact Tool Box. Version 1.1, 2007. Tool Box for Impact Monitoring and Evaluation of Savings & Credit Programme. Cochlin <http://www.ngo-ideas.net/toolbox.html>? (1.12. 2008).

Outcome Mapping (OM)

BETTINA SCHMIDT

1. Allgemeine Informationen

Outcome Mapping (OM) wurde von der Evaluierungsabteilung des International Development Research Centre (IDRC) in Ottawa, Kanada, entwickelt und 2001 veröffentlicht.

Im Rahmen einer Untersuchung stellten einige kanadische Experten gemeinsam mit Partnern aus Entwicklungsvorhaben fest, dass viele Organisationen sich mit den folgenden Problemen und Herausforderungen auseinandersetzen müssen:

- Sie müssen immer komplexere Vorgänge bewältigen;
- sie haben Probleme, nachzuweisen, wie Programme grundlegende und langfristige positive Änderungen im Leben der Zielgruppe bewirken und
- sie benötigen Instrumente zur Erhebung und Messung der Resultate (*results*), der kurz- und mittelfristigen (*outcomes*) sowie der langfristigen (*impact*) entwicklungspolitischen Wirkungen.

Ziel von OM ist, Organisationen beim Erfassen (*mapping*) der kurz- und mittelfristigen Wirkungen von spezifischen Entwicklungsprojekten und -programmen zu unterstützen.

Neu an OM ist, dass nicht die Ergebnisse und erreichten Oberziele der Maßnahmen und Programme (z. B. politische Relevanz, Armutsminderung, Konfliktreduzierung) im Mittelpunkt stehen und beurteilt werden. Im Fokus stehen Verhaltensänderungen, Lernprozesse, Beziehungen, Handlungen und Aktivitäten der Menschen und Organisationen, die direkt am Entwicklungsprogramm beteiligt sind.

Outcome Mapping kann auch als Qualitätsmanagement verstanden werden: als partizipativ und flexibel anpassbare Methode wird ein Beobachtungs- und Beurteilungsrahmen geschaffen, welcher eine Lernkultur und Lernprozesse innerhalb von Projekten, Programmen und Organisationen fördert. OM wurde entwickelt, um die Grenzen der Logical Framework Analysis (LFA), der Zielorientierten Projektplanung (ZOPP) und des Results Based Management (RBM) zu überwinden.

2. Anwendungsbereiche

Outcome Mapping kann derart angepasst werden, dass es auf Projekt-, Programm- und Organisationsebene als Monitoringsystem eingesetzt werden kann. Es kann ebenfalls für die Evaluierung laufender und abgeschlossener Aktivitäten zum Einsatz kommen.

Empfehlenswert ist, OM als Monitoring- und Evaluierungsverfahren (M&E) in der Planungsphase als Teil des allgemeinen Projektmanagements einzuführen. Somit ist es mit den verschiedenen Maßnahmen und Aktivitäten verknüpft und integraler Bestandteil des Projektmanagements. Es verbindet Prozess- und Wirkungsevaluierung und ist somit hervorragend geeignet, die komplexen Funktionsweisen und langfristigen Aspekte internationaler Entwicklungsvorhaben zu erfassen, vor allem dann, wenn kurz-, mittel- und langfristige entwicklungspolitische Wirkungen (*outcomes* und *impact*) schwierig zu erheben und zu beurteilen sind.

Es ist insbesondere für das Monitoring und die Evaluierung von Entwicklungsvorhaben geeignet, deren Resultate und Erfolge nicht allein mit quantitativen Indikatoren erfasst werden können, sondern ein umfassenderes Verständnis qualitativer und kontextbezogener Geschichten (*stories*) von Entwicklungsprozessen erfordern.

3. Analyseansatz

Outcome Mapping bietet Entwicklungsvorhaben die erforderlichen Instrumente für eine systematische und strategische Herangehensweise an das Erreichen von Resultaten. Zwei Aspekte sind wesentlich für OM:

1. Bei Entwicklung geht es primär um Menschen bzw. die Beziehungen zwischen Menschen und ihrer Umwelt.
2. Entwicklung erfolgt durch die Verhaltensänderung der beteiligten Personen.

Daher ist der Fokus von Outcome Mapping auf Personen und ihre Organisationen gerichtet. Die Handlungen und Aktivitäten der Personen rücken in den Vordergrund; mit anderen Worten: es sind die Verhaltensänderungen, die durch Monitoring und Evaluierung erfasst werden.

OM ist ein integrierter Ansatz zur Planung, zum Monitoring und zur Evaluierung, der die Komplexität der Entwicklungsprobleme berücksichtigt. Das Umfeld, in dem die meisten Projekte, Programme und Organisationen eingebettet sind, ändert sich häufig aufgrund von Faktoren, die außerhalb des Einflussbereichs und der Kontrolle des Projektteams liegen. Rahmenbedingungen, die man zu Beginn eines Projekts für selbstverständlich gehalten

hat, können sich während der Durchführung ändern. Aspekte, die zu einem bestimmten Zeitpunkt ausschlaggebend waren, können zu einem späteren Zeitpunkt an Relevanz verlieren. Daher wird ein Vorgehen erforderlich, mit dem ein Projektteam oder die Mitglieder einer Organisation schnell auf die sich ändernden Umstände reagieren können. Dies wird durch Partizipation ermöglicht. Outcome Mapping setzt auf einen partizipativen Prozess, der darauf ausgerichtet ist, im Rahmen eines Projekts, eines Programms oder einer Organisation eine Lernkultur zu schaffen, die reflektierendes und evaluatives Denken (*evaluative thinking*) fördert.

Verhaltensänderungen von Personen, Gruppen und Organisationen, die direkt am Programm beteiligt sind, werden als *outcomes* bezeichnet. Durch OM können Entwicklungsprogramme ihren Beitrag zu den Erfolgen kurz- und mittelfristiger Wirkungen (*outcomes*) erfassen, anstatt die Erfolge langfristigen Wirkungen (*impact*) zuzurechnen. Auch wenn diese kurz- und mittelfristigen Wirkungen (*outcomes*) wiederum die Möglichkeit einer langfristigen Wirkung erhöhen, ist nicht notwendigerweise eine direkte Ursache und Wirkung gegeben. Anstatt zu versuchen, die langfristigen Wirkungen der Programmpartner auf die Entwicklung zu messen, konzentriert sich OM auf das Monitoring und die Evaluierung ihrer Resultate im Hinblick auf den Einfluss den das Programm auf die Rollen, die diese Partner bei der Entwicklung spielen, hat.

Damit weicht OM grundlegend von den gebräuchlichen Ansätzen, wie zum Beispiel LFA und RBM ab, die die Beziehungen von Ursache und Wirkung zu sehr vereinfachen.

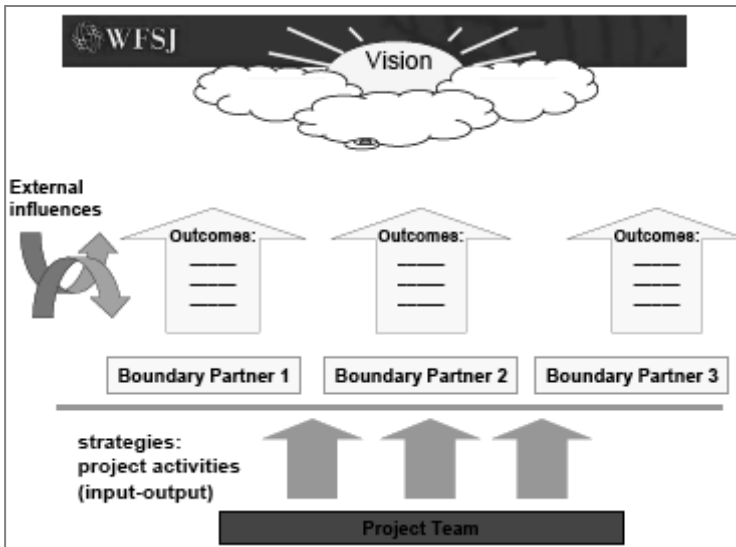
OM geht über die Erzeugung und Weitergabe von Informationen hinaus. Es berücksichtigt die aktive Beteiligung der verschiedenen Entwicklungsakteure. „Outcome Mapping umfasst nicht nur wesentliche Grundlagen für die Erstellung eines Evaluationsrahmens, sondern unterstützt Lernprozesse, erhöht die Effektivität und betont, dass Achtsamkeit auf dem Weg zum Ziel ebenso wichtig ist wie das Ankommen.“ (Michael Quinn Patton, in Earl 2001, S. ix)

Die „Projektpartner“ werden belohnt und profitieren wenn sie ihre Rollen im Entwicklungsprozess erfüllen. Entwicklungsprogramme werden honoriert für ihren Beitrag zum Entwicklungsprozess. Mit OM legen die Verantwortlichen der Entwicklungsprogramme fest, mit welchen Projektpartnern sie zusammenarbeiten. Sie entwickeln Strategien, um ihre Projektpartner mit den erforderlichen Instrumenten, Techniken und Ressourcen auszustatten, damit diese entsprechend zum Entwicklungsprozess beitragen können.

4. Durchführung

Um die Struktur von Outcome Mapping zu verstehen, setzt dies eine Erläuterung der drei Hauptbegriffe von OM voraus. Dies sind: „Projektpartner“ (*boundary partners*), kurz und mittelfristige Wirkungen (*outcomes*) und Fortschrittsmerkmale (*progress markers*).

- Projektpartner (*boundary partners*): Dies sind Personen, Gruppen und Organisationen, mit denen die Verantwortlichen des Entwicklungsprogramms direkt zusammenarbeiten. Es sind die Akteure, die Veränderungen bewirken. Wichtig ist, mit ihnen etwaige externe Faktoren, die den Prozess positiv oder negativ beeinflussen könnten, zu besprechen, um rechtzeitig erforderliche Änderungen vorzunehmen.



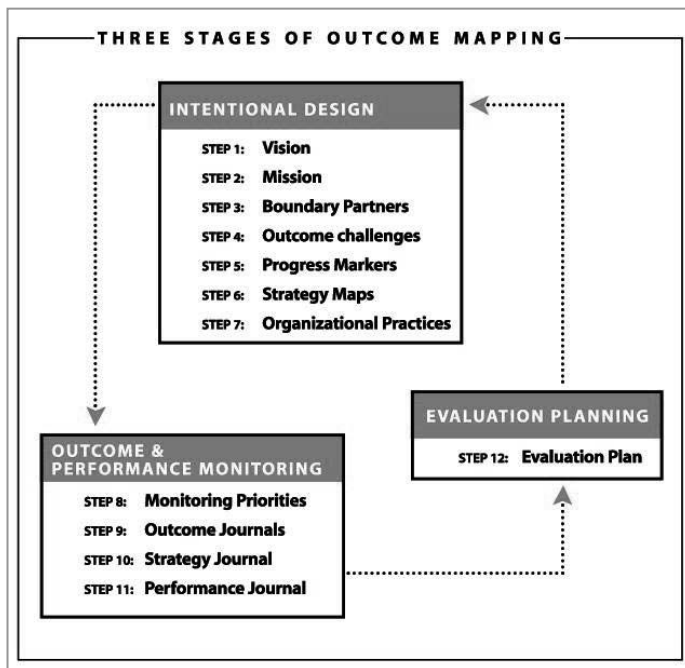
- Kurz- und mittelfristige Wirkungen (*outcomes*): Veränderungen in den Beziehungen, Aktivitäten, Handlungen oder im Verhalten von Projektpartnern (*boundary partners*), die konkret auf die Aktivitäten zurückgeführt werden können, ohne notwendigerweise direkt durch diese verursacht zu sein. Diese Änderungen zielen darauf ab, zu speziellen Aspekten der Lebensqualität, zu menschlichem und ökologischem Wohlbefinden beizutragen. Um zum Entwicklungsprozess beitragen zu können, werden Projektpartnern die erforderlichen Instrumente, Techniken und Mittel zur Verfügung gestellt.

- Fortschrittsmerkmale (*progress markers*) sind eine Reihe gestaffelter Indikatoren zu den geänderten Verhaltensweisen eines Projektpartners, die den Grad oder die Qualität der Veränderung messen.

a) Verfahrensablauf

Der vollständige OM-Prozess besteht aus drei Phasen:

In der ersten Phase, der Planungsphase (*intentional design*), stellen die Mitglieder eines Projektteams oder einer Organisation einen Konsensus zu geplanten Änderungen auf Makroebene her. Sie erarbeiten die Strategie, mit der das Ziel erreicht werden kann.

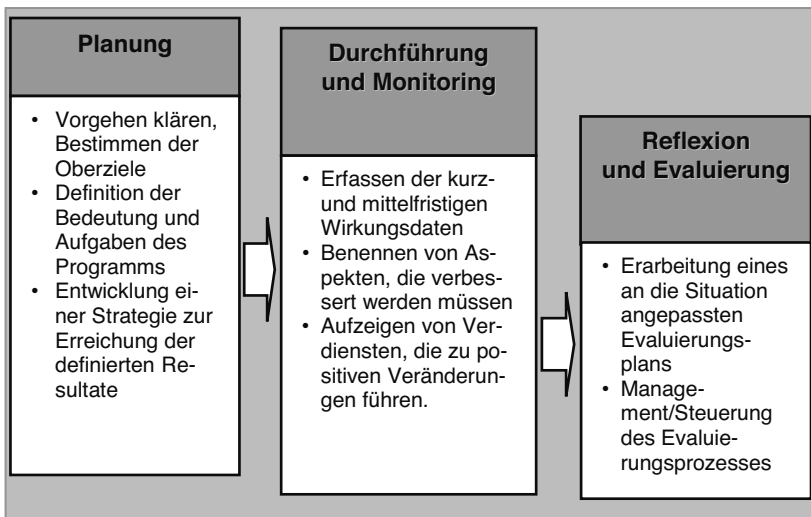


Es handelt sich dabei um einige gestaffelte Indikatoren für erforderliche Verhaltensänderungen, die für die Festlegung der Ziele mit den Projektpartnern und für das Monitoring der Ergebnisse erforderlich sind.

Phase zwei, das Wirkungs- und Umsetzungsmonitoring (*outcome & performance monitoring*), umfasst drei wesentliche Dokumente. Das Veränderungsjournal und das Strategiejournal dienen dem Monitoring der Strategien und Aktivitäten. Das Durchführungsjournal (*performance journal*)

dokumentiert die Projektdurchführung. Das Durchführungsmonitoring bietet die Möglichkeit und die Tools, um das Erreichte zu reflektieren und ebenso Verbesserungen bei der Umsetzung vorzunehmen. Hierzu werden Daten zu den Resultaten der *boundary partners* gemeinsam mit ihnen erfasst.

Die dritte Phase, die Evaluierungsplanung (*evaluation planning*), unterstützt die Arbeit durch das Formulieren von Evaluationsprioritäten. Damit wird sichergestellt, dass Mittel und Maßnahmen für die Evaluation (*evaluation resources and activities*) dort eingesetzt werden, wo sie am zweckdienlichsten sind. Das OM-Handbuch beinhaltet einen detaillierten Evaluierungsplan mit den Hauptelementen der Evaluation.



b) Akteure

Outcome Mapping erfordert die Mitwirkung der Projektmitarbeiter/-innen, der verschiedenen Projektpartner, in der Phase der Planung, des Monitorings und der Evaluation.

Da OM Wirkungen als Verhaltensänderungen definiert sind, müssen das Monitoring und die Evaluierung dies widerspiegeln. Veränderungen werden nicht nur von den Partnern realisiert: die letztendliche Verantwortung für Veränderungen tragen alle gemeinsam, die Verantwortlichen der Entwicklungsmaßnahme, Projektpartner (*boundary partners*), ihre je-

weiligen Partner und andere für das Vorhaben identifizierten Akteure. Daher sind sie auch alle an dem Monitoring und der Evaluierung beteiligt.

c) Datenerhebung

In allen drei Phasen des OM werden Daten erhoben (zwölf Schritte). In der ersten Phase werden Fortschrittsindikatoren, also Messgrößen, bezüglich der folgenden vier Fragen dokumentiert: Warum? (Vision), Wie? (Mission, Strategiepläne, *strategy map*, Organisation/Projektmanagement), Wer? (Projektpartner), und Was? (Herausforderungen bei der Erreichung kurz- und mittelfristiger Wirkungen, Fortschrittsindikatoren). Dies wird im *outcome journal* (Wirkungsjournal) dokumentiert.

In der zweiten Phase werden Strategien und Aktivitäten im Strategiejournal (*strategy journal*) festgehalten. Das *performance journal* (Durchführungsjournal) dokumentiert die erzielten Ergebnisse, die Effektivität und Effizienz der Aktivitäten.

d) Datenauswertung

Das OM-Handbuch bietet Instrumente und Arbeitsblätter für die Datenerhebung in jeder Phase des OM. Diese können einzeln verwendet und ebenso auf den spezifischen Bedarf angepasst werden. Sie sind als Hilfestellung und Arbeitserleichterung für Organisationen und Anwender gedacht, um Informationen über erzielte kurz- und mittelfristigen Wirkungen zu sammeln, zu organisieren, zu dokumentieren und zu kommunizieren.

5. Nutzung der Ergebnisse

Die detailliert ausgearbeiteten Vorlagen für die Datenerhebung und die Durchführung der Workshops setzen einheitliche Standards und sichern somit die Qualität der Datenerhebung und Partizipation der Akteure eines Vorhabens. Dies ermöglicht eine Vergleichbarkeit und schafft Transparenz. Der Lerneffekt wird erreicht, indem alle Mitarbeiter/-innen des Programms sowie Partner im gesamten Verlauf und je nach Erfordernis in Planung, Monitoring und Evaluierung einbezogen sind.

6. Anforderungen

Outcome Mapping möchte auf einfache Art und Weise, ohne Fachbegriffe und auf der Grundlage allgemeiner Mindestanforderungen an das Projekt-

management, Antworten auf komplexe Probleme und Fragen geben. Für erfahrene Evaluatoren/-innen reichen das veröffentlichte Handbuch und die Fallbeispiele in verschiedenen Sprachen (Englisch, Französisch, Spanisch) aus, um OM anzuwenden. Die erstellten Unterlagen enthalten ausführliche Informationen zur Konzeption der Workshops, Arbeitsblätter, Checklisten und weitere Materialien sowie Informationen über Dauer und die erforderlichen finanziellen Mittel.

Im Internet tauscht die OM-Community Erfahrungen in der Anwendung von Outcome Mapping aus. IDRC und einige Partnerorganisationen bieten außerdem Schulungen an. Da die Logical Framework Analysis (LFA) weit verbreitet ist, müssen Anwender/-innen von OM die Unterschiede zwischen den beiden Systemen kennen. Sie müssen auch Kenntnisse über die Anwendungen und Techniken von OM haben, die in verschiedenen Kontexten zur Anwendung gekommen sind, also praxiserprobt sind und sich somit als nützlich erwiesen haben.

7. Besondere Eignung und Grenzen des Verfahrens

Die Stärken von Outcome Mapping sind:

- Es geht um Verbesserung der Interaktion, (Re)aktion und Teilnahme in komplexen Situationen, die es zu verändern gilt.
- Es schafft eine Atmosphäre für positives, kreatives und kritisches Denken.
- Es gewährleistet die Partizipation der Teilnehmenden.
- Neben dem Monitoring von Fortschrittsindikatoren ist OM ein nützliches Management-Tool für die Strategieentwicklung.
- Es ist ein integraler Bestandteil des gesamten Management-Prozesses von der Planung über die Durchführung bis zur Erreichung der gesteckten Ziele. Es dient ebenfalls dem Qualitätsmanagement.
- Die Monitoringberichte, Interviews und Diskussionen sind Grundlage für den Lernprozess der Mitglieder des Programms und der Organisation und garantieren den Erfolg des Programms.
- Es ist besonders geeignet für das Monitoring und die Evaluierung von Entwicklungsprogrammen, deren Resultate und Erfolge nicht allein mit quantitativen Indikatoren erfasst werden können, sondern ein tieferes Verständnis des qualitativen und kontextualisierten Entwicklungsprozesses erfordern.
- Es geht um die Verwendung der Daten zur Verbesserung von Entwicklungsvorhaben, d. h. für die Erreichung der Veränderungen.

- Es kann mit anderen Methoden kombiniert werden, auch mit der LFA (ungeachtet der Unterschiede zwischen OM und LFA).
- Fachbegriffe werden vermieden und es ist für alle Beteiligten leicht verständlich (dies ist bei ZOPP und LFA nicht der Fall).

Die Einführung von Outcome Mapping kann als zeitaufwändig erlebt werden. Auch ist es nicht einfach, die Diskussionen über längere Zeit hin aufrecht und strukturiert zu halten. Dies gilt allerdings auch für andere partizipative M&E Verfahren. Outcome Mapping ist für Vorhaben ungeeignet, bei denen die Partizipation und die Meinungen der am Vorhaben beteiligten Personen nicht willkommen sind bzw. auf Grund z. B. politischer Rahmenbedingungen, Geschlechterdiskriminierung etc. freie Meinungsäußerung für alle bzw. eine Gruppe von Personen massiv eingeschränkt ist.

Obwohl Outcome Mapping die Messgrößen auf kurz- und mittelfristige Wirkungen beschränkt (die Verhaltensänderungen der Projektpartner), schafft es gleichzeitig eine Vision zur langfristigen Verbesserung des sozialen und ökologischen Umfelds, zu dem das Programm oder die Organisation betragen möchte.

Literaturhinweise

Earl, Sarah et al. (2001): Outcome Mapping. Building Learning and Reflection into Development Programs. International Development Research Centre (IDRC). www.idrc.ca.

Weblinks:

International Development Research Centre, Evaluation Unit. IDRC: www.idrc.ca.

Outcome Mapping online community: www.outcomemapping.ca.

Measuring Change. Planning, Monitoring, Evaluation in Media Development. 2007. http://www.cameco.org/files/measuring_change_1.pdf.

Participatory Impact Monitoring (PIM) / FAKT

EBERHARD GOHL

1. Allgemeine Informationen

PIM / FAKT wurde zu Beginn der 90-er Jahre durch FAKT Beratung für Management, Bildung und Technologien GmbH entwickelt, im Auftrag des für Angepasste Technologien zuständigen GTZ-Vorhabens „German Appropriate Technology and Ecoefficiency Programme“ (GATE). Es wurde den beteiligten NRO als Alternative angeboten zu dem in den GTZ-Projekten geforderten Monitoring anhand der Zielorientierten Projektplanung (ZOPP), eine GTZ-spezifische Erweiterung des Logical Framework. In jüngerer Zeit haben z. B. allein in Kenia u. a. Support for Tropical Initiatives in Poverty Alleviation (STIPA), das Arid Lands Resource Management Project (ALRMP) und das Western Kenya Community Driven Development Project, in Somalia z. B. UNICEF Somalia, PIM / FAKT eingeführt.

2. Anwendungsbereiche

PIM / FAKT wurde für die Anwendung in allen wichtigen Sektoren entwickelt, dabei besonders für die Selbsthilfeförderung, und zwar in Projekten, die an der Förderung von Gemeinschaften / Gruppen / Basisorganisationen ansetzen.

Mit PIM / FAKT werden vor allem Wirkungen auf der Mikroebene (Wirkungen der Gruppen und NRO auf lokaler Ebene) erfasst. Diese werden aus der Perspektive von drei Kategorien von Akteuren betrachtet: Gruppe (Basisorganisation); NRO oder Projekt durchführende Organisation; finanzierende Organisation.

PIM / FAKT dient dem Monitoring und der Evaluierung, kann aber in modifizierter Form auch für die Planung (ex-ante) eingesetzt werden. Ausgangspunkt ist die Erfassung von Zustandsveränderungen; in der Praxis führt dies häufig zu der Reflexion von Prozessen der Selbstorganisation.

3. Analyseansatz

Erfasst werden das Eintreten von erwarteten Wirkungen und die Vermeidung von befürchteten Wirkungen. Die Abfrage von Erwartungen und Befürchtungen geschieht unabhängig von den Planungsdokumenten. Ansatzpunkt ist jegliche relevante Veränderung, die beobachtet wird; von dort ausgehend wird analysiert, was oder wer zu dieser Veränderung beigetragen hat. Konkret: Welchen Anteil hatte die Gruppe? Welchen Anteil hatte die NRO? Welchen Anteil hatten andere Akteure / Einflüsse?

Der Umgang mit dem Zuordnungsproblem erfolgt in zwei Schritten. Im ersten Schritt handelt es sich um einen impliziten Umgang: alle aus der Sicht der Beteiligten relevanten oder signifikanten Veränderungen, d. h. die subjektiven Einschätzungen der jeweiligen Akteure (z. B. der Gruppe) werden erfasst. Im zweiten Schritt wird untersucht, inwieweit die Gruppe selbst, die NRO oder andere Akteure zu der jeweiligen Veränderung beigetragen haben. Dabei werden die eigenen Einschätzungen mit denen aus anderen Perspektiven verglichen und analysiert (z. B. der NRO). Dies sind die Joint Reflection Workshops.

4. Durchführung

a) Verfahrensablauf

Die Einführung und Durchführung von PIM / FAKT umfasst sieben Schritte (siehe auch Abbildung 12 auf der folgenden Seite).

A. Einführung vom PIM / FAKT

Vorbereitender Schritt: Analyse der Kontexte

- Analyse der Situation, z. B. Analyse des Problems, das angegangen werden soll, und Stärken und Schwächen der Gruppe
- Analyse anderer Akteure im Projektumfeld
- Austausch über die Selbstanalyse mit anderen Akteuren

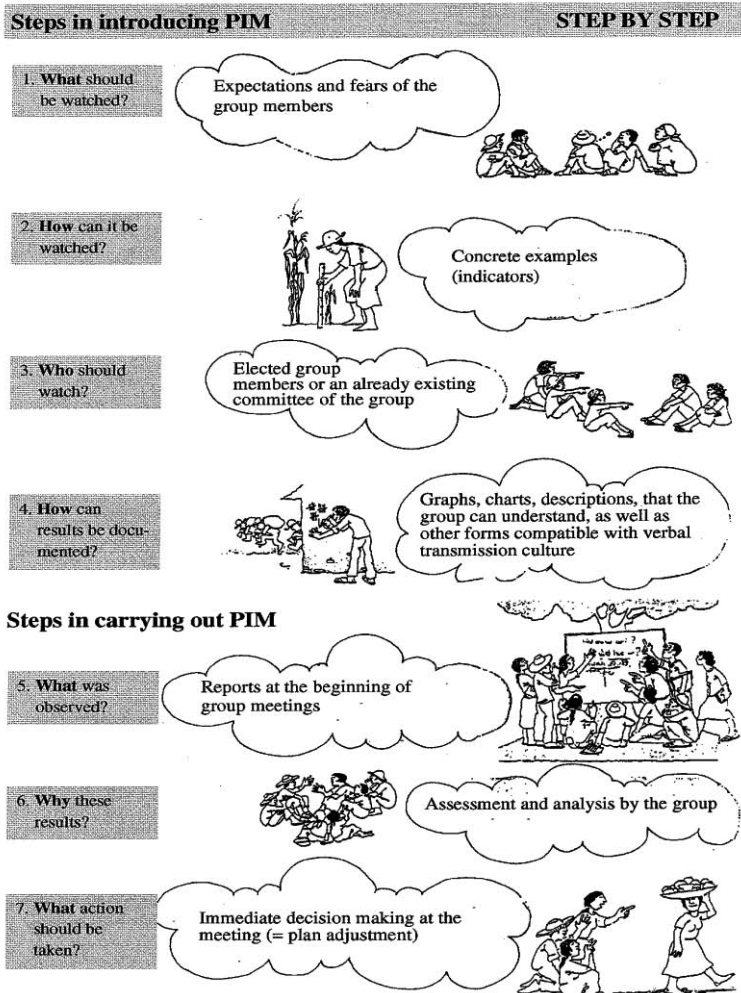
Schritt 1: Erwartungen und Befürchtung der Gruppenmitglieder ermitteln

- „Was erwarten wir von einem Projekt?“
- „Was befürchten wir von einem Projekt?“ („Was wollen wir vermeiden?“)

Sowohl aus den Erwartungen als auch aus den Befürchtungen werden Ziele identifiziert. Während von „Erwartungen“ eher technische und wirtschaft-

liche Ziele abgeleitet werden, weisen die „Befürchtungen“ häufig auf Ziele der sozialen und personalen Kompetenzen hin, die wichtige Voraussetzungen für Entwicklung sind. Von den genannten Zielen sollen für PIM / FAKT nur etwa drei bis fünf priorisiert werden.

ABBILDUNG 10: Schritte zur Einführung und Durchführung von PIM / FAKT



Schritt 2: Bildung von Indikatoren

Für die priorisierten Ziele wird ermittelt, wie die Betroffenen am besten Veränderungen in diesen Bereichen erkennen können. Es werden vier einfache Messverfahren vorgeschlagen, und Möglichkeiten der Informationssammlung erkundet.

Schritt 3: Verantwortliche Personen benennen

Es werden die Personen benannt, die die regelmäßige Beobachtung durchführen werden. Es wird vorgeschlagen, dazu auch Mini-Teams zu bilden, und diese auch bewusst aus unterschiedlichen Perspektiven beobachten zu lassen.

Schritt 4: Dokumentation und Berichterstattung über beobachtete Veränderungen vereinbaren

Es wird die schriftliche Dokumentation der beobachteten Veränderungen vorgeschlagen. Gleichzeitig werden Beispiele gezeigt, wie die Dokumentation auch öffentlich erfolgen kann, und dass dies bereits Wirkungen auslösen kann. Gleichzeitig wird verdeutlicht, dass nicht alle Informationen breit veröffentlicht werden müssen, sondern dass manche Informationen vertraulich sind.

B. Durchführung von PIM / FAKT

Schritt 5: Über beobachtete Veränderungen berichten

Die Berichterstattung über die beobachteten Veränderungen knüpft an der Möglichkeit der öffentlichen Dokumentation an. Die mündliche Berichterstattung ist jedoch zentral, in Verbindung mit der Frage nach Bestätigung oder Ergänzung der Beobachtungen (Validierung).

Schritt 6: Ursachen der Veränderungen analysieren

Mit der Frage „Warum haben wir diese Ergebnisse?“ wird zunächst in Erinnerung gerufen, was zwischendurch geschehen ist. Anschließend wird rückblickend gefragt:

- Was wurde durch die Gruppe beeinflusst?
- Was wurde durch die NRO beeinflusst?
- Was wurde durch andere Akteure beeinflusst?

Dieselben Fragen werden auch hinsichtlich ihrer Relevanz für die Zukunft gestellt:

- Was kann durch die Gruppe beeinflusst werden?
- Was kann durch die NRO beeinflusst werden?
- Was kann durch andere Akteure beeinflusst werden?

Schritt 7: Über weitere Durchführung entscheiden

Ausgehend von der zentralen Analyse in Schritt 6: „Was können wir selbst beeinflussen?“ kann die Gruppe Entscheidungen für ihre Projektdurchführung beschließen. Diese Entscheidungen sollen auch dokumentiert und den anderen Betroffenen kommuniziert werden.

Es wird empfohlen, nach einer gewissen Zeit zu überprüfen, was am Monitoring der Gruppe weiter verbessert werden könnte.

b) Akteure

Partizipation wird hier explizit verstanden als Dialog und Zusammenarbeit autonomer Akteure, die immer verschiedene „Projekte“ haben werden, mit dem Ziel der kontinuierlichen Befähigung (*empowerment*) der (benachteiligten) Menschen zur Selbsthilfe. (Das war seinerzeit nicht so üblich, damals war Partizipation eher fokussiert auf „Beteiligung an Konsentscheidungen“.)

PIM / FAKT ist sehr prozessorientiert (Betonung von *action-reflection-action* und Dialog). Damit ist es zugleich handlungsorientiert („Monitoring means continuous observation, reflection and correction of activities.“). Es handelt sich bei PIM / FAKT um eine Form von Selbstevaluierung, verbunden mit Dialog und Reflexion über die Fremdwahrnehmung.

c) Datenerhebung

Die Wirkungsbeobachtung und Datensammlung erfolgt durch die Protagonisten (NRO, Gruppe). Zum *sampling* gibt es keine Vorgaben; PIM ist konzipiert für die Selbststeuerung jeder einzelnen Gruppe.

Bezüglich der Monitoringdaten gilt:

- Zur Qualität der Indikatoren gibt es keine Vorgaben; sie sollen durch die jeweilige Gruppe gewählt und gehandhabt werden.
- Zur Messung / Beschreibung von Wirkungen werden verschiedene Möglichkeiten aufgezeigt.
- Tendaussagen sind tendenziell wichtiger als die genaue Messung.
- Der vertrauliche Umgang mit der Information wird mehrfach betont.

Bezüglich der Datenerhebungsinstrumente gibt es keine bestimmten Vorgaben, stattdessen wird bei PIM / FAKT zur Kreativität ermutigt. Die Entwicklung des Konzepts, der Berichterstattung und Auswertung erfolgt grundsätzlich in den Versammlungen der jeweiligen Gruppe.

Die Wahrnehmung und Analyse der Veränderungen aus verschiedenen Perspektiven (*triangulation*) ist expliziter Bestandteil von PIM / FAKT, in

der Form von Dialog. Indem die beteiligten Akteure angeleitet werden, ihr eigenes autonomes Wirkungsmonitoring durchzuführen, entstehen separate, auf einander abgestimmte Stränge für jede Gruppe von Akteuren. Dies drückt sich aus in abgestuften Anleitungen für PIM / FAKT für die Gruppe und die NRO (das ursprüngliche geplante Heft für die finanzierenden Organisationen wurde nicht fertig gestellt, kann aber logisch hinzugefügt werden).

d) Datenauswertung

Die Auswertung wird im Handbuch relativ ausführlich, mit vielen Beispielfragen beschrieben. Für die Joint Reflection Workshops werden die folgenden Fragen empfohlen, für die jeweils eine Reihe vertiefender Fragen vorgeschlagen werden:

1. Was hat sich verändert?
2. Was haben die (beteiligten) Menschen gelernt?
3. Welche Schritte müssen jetzt unternommen werden?
4. Wie können wir unser Wirkungsmonitoring verbessern?
5. (Für die NRO:) Was bedeutet das für unsere künftige Arbeit?

Es gibt kein festgelegtes Raster für die Datenauswertung. Erwähnt sei, dass PIM / FAKT sich explizit auf die PRA-Ansätze sowie auf Moderation (*facilitation*) bezieht.

Die Analyse der Veränderungen aus verschiedenen Perspektiven (Triangulation) ist expliziter Bestandteil von PIM / FAKT, in der Form von Dialog. Daher gibt es auch separate, auf einander abgestimmte Stränge im PIM / FAKT für jede Gruppe von Akteuren, und die Joint Reflection Workshops.

Die kritische Hinterfragung der eigenen Wahrnehmung ist unverzichtbarer Bestandteil von PIM/FAKT. Trotzdem findet bei PIM/FAKT nur eine einfache Form der Validierung oder Objektivierung statt (zwischen den Beteiligten, nicht aus einer explizit unabhängigen Perspektive). Das Hinzuziehen weiterer Perspektiven bietet sich jedoch an.

5. Nutzung der Ergebnisse

Die Ergebnisse sind als subjektive Wahrnehmungen der Beteiligten, die gemeinsam reflektiert wurden, transparent. Die Verständlichkeit ist in der Praxis sehr unterschiedlich, da keine Vorgaben für die Konkretionsebene oder die graphische Darstellung gemacht werden.

Quantifizierungen (genaue Messungen; Skalierungen; Auszählungen von Ja- und Nein-Antworten) werden im Handbuch vorgeschlagen, sind aber nicht vorgeschrieben. Das Ampel-Konzept wird als Möglichkeit dargestellt.

Die Nutzung der Ergebnisse und die Handlungsorientierung sind unmittelbarer Bestandteil von PIM / FAKT. Im Verfahren angelegte Nutzer und Funktionen ergeben sich aus der Zusammenarbeit von Gruppe und NRO, mit gegenseitiger Berichterstattung (die aber nicht genauer beschrieben ist).

6. Anforderungen

Der Zeitaufwand für diese Art von Selbststeuerung kann sehr stark variieren. Es gelten folgende Richtwerte:

- Einführung in der NRO: 3 Tage;
- Einführung in der Gruppe: 1 Tag;
- Arbeit der Beobachter der Gruppe: ca. 1 Tag pro Monat pro Person.

Die Durchführung kann für NRO und Gruppe zum Bestandteil der regulären Arbeit werden und erfordert keinen nennenswerten Extraaufwand.

Eine zentrale Voraussetzung für die Durchführung von PIM / FAKT ist, dass die NRO mit partizipativen Ansätzen arbeiten und das Konzept der Moderation praktizieren muss. Die NRO muss zudem regelmäßig (ein- bis zweimal jährlich) zu den gemeinsamen Auswertungs-Workshops einladen.

7. Besondere Eignung und Grenzen des Verfahrens

Es sind insbesondere drei Aspekte von PIM / FAKT, die es stark von anderen Verfahren unterscheidet:

1. Die bildhafte Visualisierung hat stark zur Verbreitung von PIM / FAKT beigetragen. Sie hat es plastischer, praxisorientierter, anschaulicher und weniger theoretisch erscheinen lassen.
2. Die eingebaute Flexibilität, die durch die Ausrichtung auf die Erwartungen und Befürchtungen der Beteiligten gegeben ist, ermöglicht laufende Anpassungen an das jeweils Wichtige.
3. Die Einfachheit, mit Bildern und Fragen zu arbeiten und damit Monitoring und Evaluierung auf seine Grundelemente zu reduzieren, knüpft an das natürliche Monitoring im Alltag an und ist daher leicht zu vermitteln.

PIM / FAKT ist ein Konzept, das viele Formen der Ausgestaltung zulässt. Es ist daher geeignet, erste Ideen und Ansätze von partizipativem Wirkungsmonitoring zu vermitteln. Die Grundsätze erlauben eine freie Ausgestaltung, auch unabhängig von vorgegebenen Planungs- und Monitoringdokumenten. Die Erkundung von Befürchtungen erleichtert auch das Monitoring von Risiken bzw. von Zielen, die außerhalb der Planungslogik liegen. Es befähigt somit die Beteiligten zu einem eigenständigen Monitoring, auch als Ergänzung zu vorgegebenen Monitoringverfahren.

Eine Begrenzung wird gelegentlich aber auch darin gesehen, dass keine genauen Vorgaben für die Durchführung gemacht werden. Von den Basisorganisationen und NRO wird ein aktives Interesse an einem gemeinsamen Monitoring und Dialog darüber vorausgesetzt, das in der Praxis oft nicht gegeben ist. Die NRO kann sich damit konfrontiert sehen, dass die Basisorganisationen ganz andere Erwartungen an das Projekt haben, als es den Planungen und Absichten der NRO entspricht.

Literaturhinweise

Germann, D. und Gohl, E. (1996): Participatory Impact Monitoring. Booklet 1: Group-Based Impact Monitoring. Booklet 2: NGO-Based Impact Monitoring. Booklet 3: Application Examples. Booklet 4: The Concept of Participatory Impact Monitoring. Hrsg.: GATE/GTZ und FAKT. Wiesbaden: Vieweg.

Weblink:

http://www.fakt-consult.de/modern/info/index_info.html.

Participatory Impact Monitoring (PIM) / SLE

CHRISTIAN BERG

1. Allgemeine Informationen

Die Grundlagen für das Verfahren, bei dem es sich um eine Weiterentwicklung von PIM / FAKT handelt, wurden im Rahmen eines Auslandsprojektes am Seminar für Ländliche Entwicklung (SLE) der Humboldt-Universität zu Berlin, in enger Zusammenarbeit mit der Welthungerhilfe und der indischen NRO MYRADA 1998 erarbeitet. In der Folge entstand zudem ein Handbuch mit zusätzlichen praktischen Anwendungsbeispielen. PIM / SLE dient seither, unterstützt durch die Welthungerhilfe und Novib, als Instrument für die periodische Wirkungserfassung im Rahmen des projektinternen Monitoring bei MYRADA u. a. indischen NRO (SCRIA, NYSASDRI, Sambandh u. a.).

In angepasster Form wurde das Verfahren darüber hinaus für die Evaluierung von aus Mitteln der Welthungerhilfe geförderten Projekten in Thailand, Myanmar und Kenia und die Entwicklung von partizipativen M&E-Systemen in IFAD-geförderten Projekten in Bangladesch, Indien und Vietnam genutzt.

2. Anwendungsbereiche

PIM / SLE ist für den Sektor Landwirtschaft / Ländliche Entwicklung i. w. S. besonders erprobt, prinzipiell aber auch in anderen Sektoren anwendbar (z. B. Ernährung, Gesundheit) und grundsätzlich für das M&E – zu Beginn, im Verlauf und am Ende einer Entwicklungsmaßnahme sowie ex-post – von zielgruppennahen Livelihood-Projekten aller Art geeignet (Ressourcenschutz, Einkommen schaffende Maßnahmen, Spar- und Kreditgruppenförderung, usw.).

Wirkungen werden mit Hilfe von PIM / SLE bislang in erster Linie auf der Zielgruppenebene (auf der Ebene von Individuen, Haushalten, Gruppen und Gemeinden) erfasst, allerdings erlaubt das Verfahren prinzipiell auch Wirkungsanalysen auf der Ebene des institutionellen Umfelds.

PIM / SLE erfasst schwerpunktmäßig Zustände (im Zeitverlauf), bildet aber über Wirkungshypothesen und die explizite Berücksichtigung externer Einflüsse und Entwicklungen auch Prozesse ab.

3. Analyseansatz

Jede von verschiedenen beteiligten Gruppen (Zielgruppen / Selbsthilfeorganisation – SHO, NRO / Durchführungsorganisation – DO, Geber) subjektiv als wichtig erachtete Veränderung (geplant oder ungeplant, positiv oder negativ) wird bei PIM / SLE betrachtet, sei es auf der Ebene der Nutzung (Verhaltensänderungen), des Nutzens oder – etwas eingeschränkt – der langfristigen (entwicklungspolitischen) Wirkung von Projektleistungen. Der Nachweis erfolgt mit Hilfe von Indikatoren über einen Vergleich der erwarteten bzw. befürchteten Veränderungen („Soll“) sowie wenn möglich des Ist-Zustandes „ohne Projekt“ mit dem Ist-Zustand zum Zeitpunkt der Erfassung.

Dem Zuordnungsproblem begegnet PIM / SLE auf zweifache Weise: Einmal durch die (qualitative) Erfassung der projektinternen und -externen Faktoren und des Grades ihres Einflusses auf die untersuchte Veränderung bzw. Variable, zum anderen durch einen Vergleichsgruppenansatz, z. B. den Vergleich von SHO mit unterschiedlich langer Förderung (frisch gegründet, 2-3 Jahre gefördert, am Ende / nach der Förderung) oder von Dörfern und Haushalten mit und ohne Projekt.

4. Durchführung

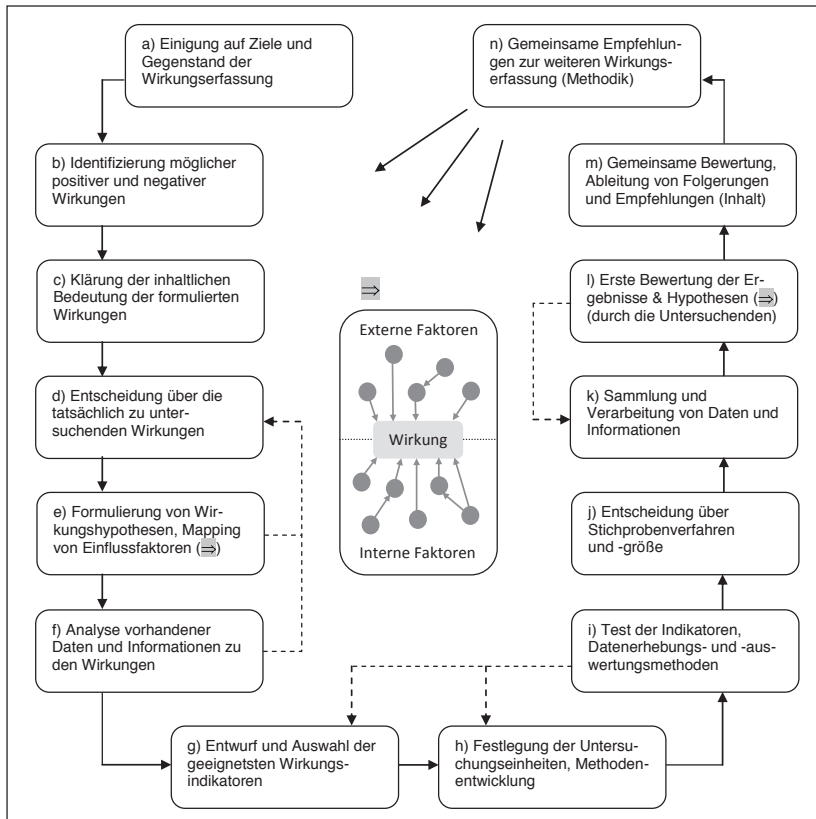
a) Verfahrensablauf

Das Verfahren ist in 20 (im Folgenden leicht zusammengefasste) Schritte gegliedert (siehe auch Abbildung 11):

- (a) Einigung zwischen den Beteiligten auf die Ziele der Wirkungserfassung – Steuerung, Lernen, Rechenschaftslegung, Lobbying und Fundraising etc. – und Entscheidung über die zu untersuchende Entwicklungsmaßnahme oder Projektkomponente, Ableitung der Konsequenzen für die folgenden Schritte;
- (b) Identifizierung möglicher Wirkungen anhand von Projektunterlagen und/oder durch offene Befragung mit der Leitfrage: Welche Erwartungen und Befürchtungen bestehen hinsichtlich Veränderungen, die das Projekt / Programm bewirkt (hat)?;

- (c) Klärung von Schlüsselbegriffen und Einigung auf die (inhaltliche) Bedeutung der formulierten Wirkungen mit der Leitfrage: Was verstehen wir unter ...?;
- (d) Entscheidung über die tatsächlich zu untersuchenden Wirkungen (Priorisierung);
- (e) Formulierung von Wirkungshypothesen, *mapping* projektinterner und -externer (positiver wie negativer) Faktoren und ihrer Bedeutung für die Wirkungen, ggf. Revision der Entscheidung über zu untersuchende Wirkungen, falls sich z. B. abzeichnet, dass die projektinternen Faktoren keine oder nur eine sehr geringe Bedeutung haben;
- (f) Analyse bereits vorhandener Informationen über die Wirkungen, ggf. Revision der Entscheidung über zu untersuchende Wirkungen, falls sich z. B. abzeichnet, dass bereits hinreichend verlässliche Informationen vorliegen;
- (g) Entwurf von möglichen Wirkungsindikatoren und kriteriengestützte Auswahl der geeignetsten Indikatoren;
- (h) Festlegung der Untersuchungseinheiten und Erarbeitung der Datenerhebungs- und -auswertungsmethoden;
- (i) Test – und ggf. Überarbeitung – der Indikatoren und Methoden;
- (j) Entscheidung über das Stichprobenverfahren;
- (k) logistische Vorbereitung der Wirkungserfassung, Sammlung und Verarbeitung von Daten und Informationen;
- (l) Analyse und vorläufige Bewertung der Ergebnisse und Wirkungshypothesen (durch die Untersuchenden), ggf. Sammlung ergänzender Informationen;
- (m) gemeinsame Bewertung der Ergebnisse und Ableitung von Schlussfolgerungen und Empfehlungen z. B. für die weitere Projektgestaltung (durch die beteiligten Akteure / Akteursgruppen);
- (n) gemeinsame Erarbeitung von Empfehlungen zur weiteren Wirkungserfassung (Methodik).

ABBILDUNG 11: Verfahrensschritte bei PIM / SLE



b) Akteure

Je nach Einsatzbereich von PIM / SLE sind die verschiedenen Stakeholder in unterschiedlichem Maße an seiner Durchführung beteiligt:

- Bei SHO-gestütztem PIM / SLE: Aktive Rolle der Zielgruppen in allen Phasen des Prozesses (beim ersten Durchlauf extern moderiert, z. B. durch NRO/DO-Mitarbeiter, und einzelne Schritte des Verfahrens vereinfacht), hohe Beteiligung der NRO/DO bei den Schritten (m) und (n) des Verfahrens.
- Bei NRO/DO-gestütztem PIM / SLE: Aktive Rolle der NRO/DO-Mitarbeiter in allen Phasen des Prozesses (beim ersten Durchlauf moderiert durch externe Fachkraft), hohe Beteiligung der Zielgruppen bei den Schritten (b), (g), (m) und (n) des Verfahrens.
- Bei (externen) Evaluierungen: Hohe Beteiligung der NRO/DO bei den Schritten (a) bis (e), (g) bis (k) sowie (m) und (n) des Verfahrens, hohe Beteiligung der Zielgruppen bei den Schritten (b), (g), (m) und (n) des Verfahrens.

Capacity building (v. a. durch *training on the job*) ist in dem Verfahren angelegt und i. d. R. notwendige Voraussetzung für die Durchführung von PIM / SLE. Es ist in erster Linie ein periodisches Monitoring- und damit Selbstevaluierungsverfahren, aber auch bei externen Evaluierungen (mit hohem Partizipationsgrad, s. o.) einsetzbar.

c) Datenerhebung

Als Datenquellen für die Wirkungserfassung dienen – in Abhängigkeit von dem jeweiligen Indikator – in erster Linie die beteiligten Stakeholder, v. a. die Zielgruppen und ihre Vertreter, daneben auch (schriftliche) Sekundärquellen und Ressourcepersonen. Die Stichprobenziehung erfolgt i. d. R. nicht durch eine strenge Zufallsauswahl, aber in jedem Fall durch ein geplantes, den konkreten Umständen angemessenes Verfahren (siehe Schritt (j) oben). Als Erhebungsinstrumente dienen

- bei der Vorbereitung der Wirkungserfassung (Schritte (a) bis (j) des Verfahrens): Brainstorming, Gruppendiskussion, Quellenstudium;
- bei der Wirkungserfassung: In Abhängigkeit von dem jeweiligen Indikator in erster Linie Einzel- und Gruppeninterview, häufig unterstützt durch Visualisierungstools (Karten, Diagramme etc.), direkte Beobachtung, direkte Messung.

Die vorhandenen Handbücher und Evaluierungsberichte enthalten *ready-made* Instrumente zur Erfassung verschiedenster Wirkungen, die ggf. für den konkreten Bedarf angepasst werden müssen.

d) Datenauswertung

Für die Datenauswertung sind keine speziellen Instrumente vorgesehen, da eine einfache statistische Auswertung (z. B. mit Hilfe von Microsoft Excel[®]) in aller Regel möglich ist. Für jeden Indikator wird ein (wo notwendig mehrstufiges) „Analyseblatt“ entwickelt, das die systematische Eingabe und Verarbeitung der gewonnen Daten / Informationen – mit oder ohne EDV – erlaubt.

Das Instrument des Joint Reflection Workshop (Schritte (m) und (n) des Verfahrens) erlaubt einen Abgleich der unterschiedlichen Einschätzungen der Beteiligten. Eine systematische Triangulation von Methoden und Quellen der Information ist in den bisherigen Anwendungen des Verfahrens nicht verankert, PIM / SLE ist allerdings prinzipiell hierfür offen (im Rahmen des Schrittes (h) des Verfahrens). Eine Einschätzung der Validität der Indikatoren ist im Verfahren vorgeschrieben. Zusätzlich erlaubt die (statistische) Auswertung sowie der Joint Reflection Workshop (s. o.) eine gewisse Validierung der Daten / Informationen.

5. Nutzung der Ergebnisse

Die Ergebnisse der Wirkungserfassung werden leicht verständlich (bei Bedarf auch rein bildlich) dargestellt und sind durch ihre Präsentation und eingehende Diskussion im Rahmen des Joint Reflection Workshop für alle Beteiligten transparent. Durch die durchgehende Bezugnahme auf die partizipativ erarbeiteten Indikatoren sind die Ergebnisse auch hinreichend konkret und nachvollziehbar für alle Beteiligten (einschließlich Zielgruppen).

Eine Quantifizierung der Analyseergebnisse (in Zahlen, Prozentsen etc.) ist durchgängig im Verfahren angelegt. Eine Aggregation der Daten ist allerdings nur bedingt möglich, da die Indikatoren lokal üblicherweise spezifisch sind.

Die am Joint Reflection Workshop beteiligten Akteure (Zielgruppenvertreter, NRO/DO-Mitarbeiter, ggf. Geber) ziehen gemeinsam Schlussfolgerungen aus den Ergebnissen und verständigen sich auf Empfehlungen hinsichtlich der weiteren Gestaltung und Durchführung sowohl des Projekts (Inhalt) als auch der Wirkungserfassung (Methodik). Ein Monitoring der Umsetzung der Empfehlungen ist nicht explizit im Verfahren angelegt,

ergibt sich aber – im Falle der Institutionalisierung von PIM / SLE – indirekt aus der regelmäßigen / periodischen Wiederholung dieses Events.

6. Anforderungen

Der Aufwand für PIM / SLE korreliert direkt mit der Zahl der zu untersuchenden Wirkungen und Indikatoren, der Stichprobengröße sowie den Personalkosten. In Berg (1998, S. 142) wurde errechnet, dass die periodische Erfassung, Auswertung und Dokumentation von sieben Wirkungen⁵² (mit insgesamt 12 Indikatoren) eines SHO-Förderprojekts in einem Subdistrikt (*taluk*) in Südindien durch lokale NRO-Mitarbeiter auf der Basis von insgesamt 330 Einzel- und Gruppeninterviews nach heutigem Wert rund 600 EUR pro Jahr für Personal, Logistik und Material kostet. Bei einer Selbstevaluierung durch Zielgruppen (z. B. Selbsthilfegruppen) wäre der Aufwand noch erheblich geringer. Allerdings sind die Kosten für die Befähigung der NRO/DO zur Durchführung von PIM / SLE nicht unerheblich, besonders wenn hierfür internationale Berater bzw. Gutachter eingesetzt werden.

Das Verfahren benötigt qualifizierte Unterstützung v. a. für das ca. 3-tägige Training von Projektmitarbeitern zu Beginn der Evaluierung oder am Beginn der Einführung von PIM / SLE als Monitoringinstrument, für die Indikatorenentwicklung, Datenverarbeitung und -analyse sowie die Moderation des Joint Reflection Workshop. Erfahrungen der NRO/DO-Mitarbeiter in Interviewführung und Anwendung von Instrumenten aus der RRA/PRA-Toolbox sind ebenfalls entscheidend.

Der Einsatz von PIM/SLE erfordert einen bereits bestehenden partizipativen Projektansatz. Eine stabile Organisationsstruktur und gute Kommunikationskultur (innerhalb der NRO/DO sowie zwischen Zielgruppen und NRO/ DO) sind wünschenswert, u. a. da die Erhebung von Daten und deren Aus- und Bewertung in Reflektionsworkshops dialogorientiert ist. Das Interesse an PIM / SLE (bzw. die Lernbereitschaft) muss in der Organisation vorhanden sein. Es ist ein gewisser Mehraufwand im Projekt, wenn PIM / SLE als Monitoringinstrument eingeführt werden soll. In diesem Fall ist ein bereits bestehendes M&E-System von Vorteil, in das das Verfahren integriert werden kann.

⁵² 1. erhöhtes Wissen und Bewusstsein der SHO-Mitglieder; 2. Entwicklung ihrer individuellen Fähigkeiten; 3. gestiegene Einkommen und Ersparnisse; 4. gestiegene Arbeitsbelastung von Frauen; 5. gestiegene Entscheidungsmacht von Frauen in der Familie sowie 6. ihr Einfluss in der Gemeinde; 7. Entwicklung von Netzwerken der SHO.

7. Besondere Eignung und Grenzen des Verfahrens

Stärken des Verfahrens liegen in der Anwendbarkeit durch lokale NRO/DO und Zielgruppen, der Nutzung aussagekräftiger, weitestgehend SMARTer Indikatoren sowie dem Wirkungsnachweis durch Vergleichsgruppen. Von Vorteil ist auch, dass die Wirkungserfassung nicht von der Qualität der Projektplanung abhängig ist.

Schwächen bestehen vor allem in dem unzureichend verankerten *cross-checking* von Daten / Informationen sowie dem nicht explizit angelegten Monitoring der Umsetzung von Empfehlungen.

Literaturhinweise

- Berg, C. et al. (1998): NGO-Based Participatory Impact Monitoring of an Integrated Rural Development Project in Holalkere Taluk, Karnataka State, India. Berlin: SLE (Schriftenreihe des Seminars für Ländliche Entwicklung Nr. S180).
- Schürmann, A. (2002): Participatory Impact Monitoring of Self-Help Groups and Watersheds. A User's Handbook. Bonn: Welthungerhilfe.

Participatory Livelihood System Analysis (PaLSA)

DIRK GUENTHER / MARKUS FIEBIGER

1. Allgemeine Informationen

Das Verfahren „Participatory Livelihood System Analysis (PaLSA)“ dient der partizipativen Planung von Projekten der Entwicklungszusammenarbeit und der Analyse der Veränderung von Lebensbedingungen (Wirkung), in Zusammenarbeit mit der Zielgruppe. Das Verfahren baut konzeptionell auf Überlegungen der Systemanalyse (Vester 2002) sowie dem MAPP-Ansatz (Method for Impact Assessment of Programmes and Projects, Neubert 2004) auf.

Das Verfahren wurde 2005 im Rahmen einer SLE-Studie entwickelt (Weingärtner et al. 2005) und von der FAO im Rahmen des Livelihood-Support-Programms in Kambodscha angewandt. Die Welthungerhilfe setzt es seit 2006 in modifizierter Form als Anwendung im qualitativen MDG-Monitoring innerhalb ihres Sektorvorhabens Millenniumsdörfer ein. Aktuell wird PaLSA im Rahmen eines Beratungseinsatzes im Zuge des philippinischen Dezentralisierungsprozesses zur Erstellung von Gemeindeentwicklungsplänen auf Mindanao verwendet.

2. Anwendungsbereiche

Das Verfahren wurde insbesondere für die Anwendung in Projekten der ländlichen Entwicklung (Armutsbekämpfung, Einkommen schaffende Maßnahmen, Förderung der Landwirtschaft, etc.) konzipiert. Die Anwendung ist auch für Projekte im Bereich der Konfliktbearbeitung denkbar, jedoch noch nicht erprobt.

Das Verfahren kann im gesamten Projektzyklus von der Situationsanalyse, der Projektidentifikation, der Planung, dem Monitoring bis zur Evaluation (Datengewinnung während der Evaluation) eingesetzt werden und sollte idealerweise projektbegleitend sein. PaLSA erfasst die wichtigsten Determinanten eines Systems (z. B. Dorfgemeinschaft oder Konflikt), welche grundlegenden Prozesse darin ablaufen und in welchem Zustand sich das System zu bestimmten Zeitpunkten befindet.

Bei der Verwendung des Verfahrens in der Planung wird ermittelt, welche Komponenten das System definieren, welche Wechselwirkungen bzw. Regelmechanismen innerhalb des Systems existieren und wo im System am erfolgversprechendsten mit Projektaktivitäten anzusetzen ist, um dieses positiv zu beeinflussen. Hierfür ist nur Phase I des Verfahrens anzuwenden (siehe Abbildung weiter unten).

Bei der Verwendung als Monitoringverfahren oder in der Evaluation wird ermittelt, ob tatsächlich ein Zusammenhang zwischen positiven Veränderungen im System (bei der Zielgruppe) und einzelnen Projektaktivitäten besteht und wie diese Wirkung ausgeprägt ist (Betrachtung von Effizienz und Effektivität einzelner Projektaktivitäten). In diesem Fall sind alle Phasen des Verfahrens anzuwenden.

3. Analyseansatz

Wirkungsketten stellen ein stark vereinfachtes Model der menschlichen Wirklichkeit dar, in der nur lineare Kausalbeziehungen abgebildet werden. Die reale Lebensumwelt wird dagegen von vielschichtigen und komplex miteinander verbundenen Faktoren und Einflüssen gebildet, die fördernd, hemmend oder neutral aufeinander wirken. Sie stellt ein komplexes System von direkten und indirekten Wechselwirkungen, Rückkopplungen und Regelkreisen dar. Mit PaLSA lassen sich die wichtigsten Komponenten eines Systems und seine inhärenten Wechselbeziehungen identifizieren. Zudem werden geplante und ungeplante, direkte und indirekte sowie leicht und schwer zuordenbare Wirkungen von Entwicklungsmaßnahmen auf das System erfasst. Dabei geht der Ansatz immer vom Kontext aus.

Das Verfahren lässt kein Zuordnungsproblem aufkommen, da es Wirkungen (Veränderungen im System) in direkte Beziehungen zu Projektaktivitäten setzt.

Das Verfahren basiert auf subjektiven (qualitativen) Einschätzungen der Zielgruppe. In der Gruppenarbeit besteht viel Raum zur Diskussion der Einschätzungen. Es steht frei, die von der Zielgruppe identifizierten Beziehungen zwischen Projektaktivitäten und Veränderungen des Systems (Wirkungen) anhand objektiver Indikatoren zu überprüfen und ggf. zu quantifizieren.

Das ursprüngliche Verfahren (siehe Strele et al., 2006a) wurde so konzipiert, dass nur abgefragt wird, ob es aufgrund einer spezifischen Aktivität zu positiver Veränderung des Systems kam und wie stark diese ausgeprägt ist. Unbeabsichtigte, negative Wirkungen werden mit dem Verfahren nicht automatisch erfasst. Dies kann jedoch erreicht werden, in dem man Schritt 6 des Verfahrens (Activity Influence Matrix, siehe Abbildung) leicht erweitert.

4. Durchführung

a) Verfahrensablauf

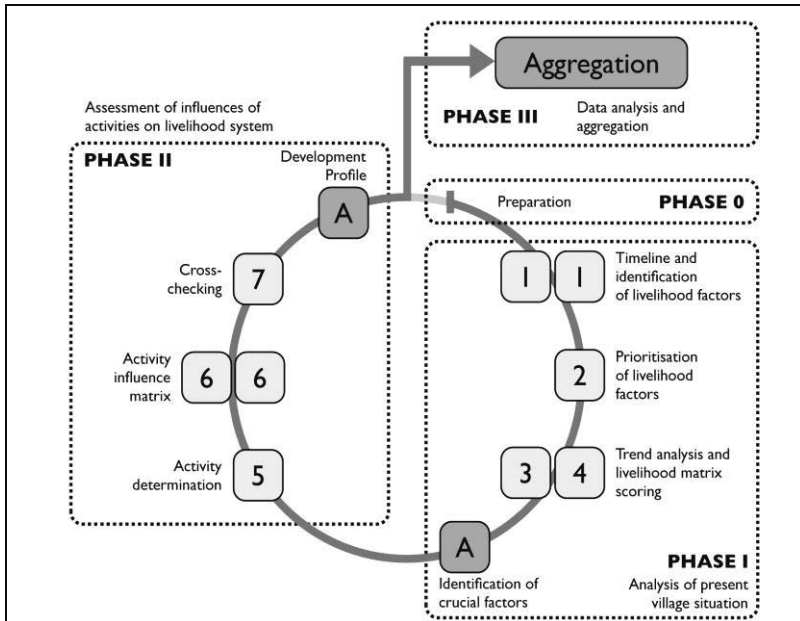
Der stark partizipative Ansatz beruht auf einer Kombination von „Werkzeugen“, die z. T. aus klassischen PRA-Methoden bestehen, und entlehnt darüber hinaus Elemente des Livelihood-Ansatzes (nach DFID), der Systemanalyse und von MAPP. Abgesehen von der Vorbereitungsphase (Phase 0) lässt sich das Verfahren in drei weitere Phasen untergliedern:

Phase I besteht aus einer Abfolge von vier partizipativen Übungen, die auf Gemeindeebene mit der Zielgruppe geführt werden. Ziel dieser Phase ist die Bestimmung derjenigen Faktoren, welche die Lebenssituation der Zielgruppen am stärksten determinieren sowie das Zusammenspiel ebendieser Faktoren. Die Identifikation dieser Faktoren folgt den Dimensionen des Livelihood-Ansatzes nach DFID (DFID 2004). Im angegliederten Analyseschritt werden diejenigen Faktoren identifiziert, welche sich am einfachsten über Projektinterventionen beeinflussen lassen und eine positive Veränderung des Gesamtsystems bewirken können (= *crucial livelihood factors*).

Die Phasen II und III dienen dem eigentlichen Wirkungsmonitoring sowie der Projektsteuerung. Phase II besteht ebenfalls aus einer Kombination von partizipativen Übungen auf Gemeindeebene und einem Analyseschritt: Nachdem die zu bewertenden Projektaktivitäten bestimmt wurden, wird an Hand einer Übung die Wirkung der ausgewählten Aktivitäten auf die in Phase I identifizierten, das System am stärksten bestimmenden Faktoren (*crucial livelihood factors*) ermittelt und die Ergebnisse mittels eines einfachen *cross-checks* überprüft. Im anschließenden Analyseschritt werden die Ergebnisse erneut ausgewertet und nach einem vorgegebenen Schema dokumentiert.

Handelt es sich bei dem durchgeführten Projekt um ein Vorhaben, dass in einem größeren regionalen Rahmen an mehreren Standorten oder in mehreren Dörfern durchgeführt wird, lassen sich in Phase III die Ergebnisse der ersten beiden Phasen aggregieren. Das Ergebnis ist ein Portfolio, das eine Übersicht über Effizienz (= *relative influence of a project activity on the system crucial livelihood factors*) und Effektivität (= *average change of livelihood situation after beginning of specific activity*) der einzelnen Aktivitäten gibt. Aufgabe der Steuerung ist es nun, die Aktivitäten auf ihr schwächeres oder stärkeres Abschneiden hin genauer zu betrachten und die Ursachen hierfür anhand der dokumentierten Ergebnisse der im Rahmen des Verfahrens durchgeführten Diskussionen mit der Zielgruppe zu analysieren und steuernd einzugreifen.

ABBILDUNG 12: Durchführungsphasen von PaLSA



Quelle: Strele et al. (2006a)

b) Akteure

PaLSA ist sowohl für Selbst- als auch für externe Evaluationen geeignet. Die Methodik ist stark partizipativ und es findet eine intensive Diskussion mit den Zielgruppen während der Teilschritte statt. Zielgruppenvertreter werden durch die Teilnahme an der Analyse geschult, Veränderungen ihrer Lebensbedingungen zukünftig auch selbst kritisch zu betrachten und systematischer zu verfolgen (*capacity building*).

c) Datenerhebung

Die Daten werden auf Workshops partizipativ mit den Zielgruppen erarbeitet, wobei Untergruppen gebildet werden können. Die Diskussionen werden in den Gruppen strukturiert geführt und die Ergebnisse visualisiert. Verschiedene Perspektiven werden berücksichtigt.

d) Datenauswertung

Es erfolgt eine halbautomatisierte Datenauswertung anhand der Dokumentation der Gruppenarbeiten und mit Hilfe einer vorformatierten Oberfläche zur Tabellenkalkulation (z. B. Microsoft Excel®). Dabei folgt die Auswertung einem festgelegten Raster. In den Auswertungsformaten besteht die Möglichkeit, Diskussionshintergründe und konträre Ansichten zu dokumentieren. Es steht frei, die Ergebnisse anhand objektiver Indikatoren zu überprüfen und ggf. zu quantifizieren.

5. Nutzung der Ergebnisse

Die Ergebnisse werden grafisch und tabellarisch dargestellt und sollten mit einem zusammenfassenden Kommentar versehen werden. In der Darstellung wird besonderer Wert auf Transparenz und Nachvollziehbarkeit gelegt. Eine Aggregation der Ergebnisse auf regionalem Niveau ist möglich, sofern es sich um eine homogene geographische Einheit handelt (z. B. Küstenebene, Gebirgstal, Flussdelta etc.).

Die Ergebnisse werden für Planung, Monitoring und Evaluation genutzt. Diese Nutzung zu verschiedenen Momenten des Projektzyklus ist im Verfahren angelegt, es werden dann jeweils andere Phasen des Verfahrens angewandt und kombiniert. Die Untersuchung kann jederzeit wiederholt werden und so die Ergebnisse fortgeschrieben werden.

6. Anforderungen

Der Aufwand für die Durchführung (Zeit und Kosten) stellt sich in etwa wie folgt dar:

Wird das Verfahren nur für die Planung eingesetzt, muss nur Phase I der Methodik durchgeführt werden. Hierfür sind pro Standort ein vier- bis fünfstündiger Workshop sowie etwa drei Stunden für die Auswertung der Daten anzusetzen. Als Grundlage für den Einsatz im Monitoring oder in der Evaluation sind pro Standort jeweils ein weiterer 4-stündiger Workshop und ca. drei Stunden für die Auswertung notwendig. Während der Laufzeit eines Projektes sollte die Monitoring-Komponente ein- bis zweimal pro Jahr wiederholt werden.

Für die Gruppendiskussionen ist eine, möglichst externe, Moderation erforderlich. Bei den üblichen Gruppengrößen von bis zu 40 Personen sind hierfür zwei Personen vorzusehen, bei Gruppen unter 15 Teilnehmern ist eine Person ausreichend. Die Moderatoren müssen in der Methodik geübt sein. Bei Neueinsteigern ist hierfür ein zwei- bis dreitägiges Training erforder-

derlich. Zur Moderation sind entweder gute Kenntnisse der Lokalsprache oder zusätzlich eine kompetente Übersetzung pro Moderator notwendig.

7. Besondere Eignung und Grenzen des Verfahrens

PaLSA ist ein qualitatives und kein quantitatives Verfahren. Es lässt vernetzte Wirkungsgefüge erkennen und lässt Sensitivitätsuntersuchungen der Faktoren (puffernd, fördernd oder kritisch, aktiv oder passiv) zu.

Stärken des Verfahrens liegen in der einfachen und transparenten Anwendung und Dokumentation. Das Verfahren ist offen und kann zur Planung, idealerweise ohne Projektfokus, eingesetzt werden. Die Einbindung des Verfahrens in ein bereits bestehendes Monitoringsystem ist ebenfalls möglich. Das Verfahren ist partizipativ und fördert die Bildung von Kapazitäten. Das Verfahren ist für komplexe Situationen geeignet. Die Ergebnisse sind nur schwer manipulierbar.

Die Anwendung erfordert allerdings gute Moderationserfahrung bzw. ein 2-3-tägiges Training in der Methodik. Auch ist die Gruppenzusammensetzung (*Gender*, Alter, Armut, etc.) für die erfolgreiche Anwendung entscheidend. Das Verfahren bereitet z. T. kleinere Schwierigkeiten für Alphabet/-innen. Eine weitere mögliche Einschränkung ist, dass das rückwärtige Erinnerungsvermögen der Zielgruppen z. T. verzerrt ist.

Im ursprünglichen Design der Methodik ist die gezielte und systematische Erfassung negativer Projektwirkungen nicht vorgesehen. Sie kann jedoch in wenigen einfachen Schritten ohne großen Aufwand in den Ablauf mit aufgenommen werden.

Literaturhinweise

- DFID (2004): *Introducing Sustainable Livelihoods. A Trainers' Guide*.
- Neubert, S. (2004): Wirkungsanalysen der entwicklungspolitischen Zusammenarbeit sind machbar. DIE. Analyse und Stellungnahmen, Nr. 4. Bonn: DIE.
- Strele, M. et al. (2006a): *Linking Programmes and Poor People's Interests to Politics Experiences from Cambodia*. FAO Livelihood Support Programme. März 2006.
- Strele, M. et al. (2006b): *Participatory Impact Monitoring for Sustainable Livelihoods*. Paper at EASY ECO Conference. Saarbrücken.
- Vester, F. (2002): *Die Kunst vernetzt zu Denken*. München: dtv.
- Weingärtner, L. et al. (2005): *Poverty and Food Security Monitoring in Cambodia*. Berlin: SLE.

Participatory Methodology for Measuring the Contribution of Volunteering to Development (MCVD)

CASPAR MERKLE

1. Allgemeine Informationen

Das Verfahren wurde 2005/2006 unter der Leitung von United Nations Volunteers (UNV) entwickelt. Der Prozess wurde von folgenden Organisationen begleitet: Canadian Crossroads International (CCI), Skillshare International (UK), Australian Volunteers International (AVI), Deutscher Entwicklungsdienst (DED), Japan Overseas Co-operation Volunteers (JOCV) und Volunteer Service Abroad (VSA) New Zealand. Derzeit wird das dazugehörige Handbuch von UNV überarbeitet und editiert, geplantes Erscheinungsdatum ist Sommer 2008 unter dem Titel „A Participatory Methodology for Measuring the Contribution of Volunteering to Development“.

2. Anwendungsbereiche

Das Verfahren zur Wirkungsanalyse von Freiwilligeneinsätzen wurde für ein breites Spektrum von Sektoren entwickelt, in denen Freiwillige einen Beitrag zur Erreichung der Millenniums-Entwicklungsziele leisten. Es kann in jeder Art von Programm oder Projekt im Rahmen der multi-/bilateralen Zusammenarbeit, auf der Mikro- und Mesoebene und bedingt auch auf der Makroebene angewandt werden. MCVD findet sowohl im Verlauf einer Entwicklungsmaßnahme Verwendung als auch am Ende, sowie ex-post. Das Verfahren wurde von UNV in 12 Ländern getestet, was zu einer ersten Version des Handbuchs führte. Im Jahr 2007 wurde es von UNV in einer Reihe von Workshops zu den folgenden Themen angewendet: Krisenbewältigung und -prävention, Postkonfliktbewältigung, Aufbau nationaler Strukturen für Freiwilligenengagement, Jugendpartizipation, sowie lokale Regierungsführung.

Der Schwerpunkt des Verfahrens liegt auf der Erfassung von Prozess- und Verhaltensänderungen, die auf Freiwilligeneinsätze zurückgeführt werden können. Beispiele für solche Veränderungen sind ein verstärktes lokales

Engagement der lokalen Bevölkerung, ein vermehrter Gebrauch partizipativer Ansätze in der Durchführung einer Maßnahme, der Aufbau eines Dialogs zwischen Regierung und lokaler Bevölkerung, etc.

3. Analyseansatz

Es werden mit Hilfe von MCVD geplante und ungeplante sowie direkte und indirekte Wirkungen nachgewiesen. Ausgangspunkt ist der „Input“, in diesem Fall der Freiwilligeneinsatz, der als Ursache von Veränderungen betrachtet wird. Konkurrierende Ursachen werden durch Fragestellungen wie „Was wäre ohne den Freiwilligeneinsatz passiert?“ berücksichtigt. Die Wirkungseinschätzung erfolgt subjektiv in der Gruppe. Das Verfahren verfolgt keinen experimentellen Ansatz, insofern wird die kontrafaktische Frage nur näherungsweise beantwortet. Die Ergebnisse aus der Gruppe werden im Anschluss durch Diskussion mit anderen Projektbeteiligten und/oder Experten validiert und trianguliert.

4. Durchführung

a) Verfahrensablauf

Das Verfahren besteht aus einer Serie von partizipativen Workshops, in denen der Beitrag von individuellen Freiwilligen in einen größeren Entwicklungszusammenhang gesetzt wird. Kernfragen bei den Workshops sind u. a. folgende:

- Was sind die Wirkungen von Freiwilligeneinsätzen?
- Wie werden diese Wirkungen erzielt?
- Was sind die *lessons learned* und Empfehlungen aus den untersuchten Freiwilligeneinsätzen?

Die Analyse basiert auf qualitativen Primärdaten von individuellen Freiwilligen, die im Laufe des Prozesses von Projektbeteiligten und/oder externen Experten auf verschiedenen Ebenen kritisch hinterfragt werden. Dabei werden die Beiträge von Freiwilligen entsprechend einer Wirkungskette (*logical framework*) aufgearbeitet und strukturiert. Gender-Aspekte sind integraler Bestandteil der Analyse. Einige Elemente von MCVD sind an andere bestehende Verfahren angelehnt, insbesondere Outcome Mapping und Most Significant Change.

Die Workshops finden auf folgenden Ebenen statt (siehe Abbildung):

- Projektebene:

Auf der Projektebene werden individuelle Beiträge von Freiwilligen anhand von Most Significant Change Stories aufgearbeitet. Die Wirkungen werden nach Wichtigkeit geordnet und anhand ihres Beitrags zu (inter-) nationalen Entwicklungszielen analysiert. Zu diesem Zeitpunkt wird auch die kontrafaktische Frage gestellt: „Was wäre ohne den Freiwilligeneinsatz passiert?“ Der Beitrag der Freiwilligen wird außerdem unter den Aspekten Alter, Geschlecht, nationaler/internationaler Freiwilligeneinsatz analysiert. Zum Abschluss des Workshops findet eine Stärken- und Schwächen- (SWOT-) Analyse statt, um eine systematische Dokumentation der Erfahrungen zu gewährleisten und für zukünftige Freiwilligeneinsätze nutzbar zu machen.

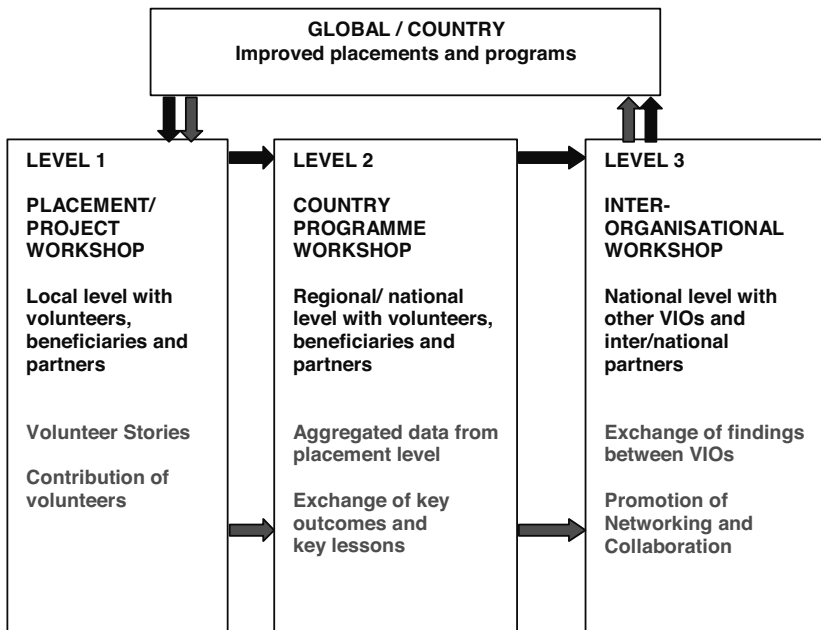
- Länderebene:

Auf der Länderebene werden die Ergebnisse und die Analyse vorangegangener Workshops weiter aggregiert und der kumulative Beitrag von Freiwilligeneinsätzen zur Erreichung (inter-)nationaler Entwicklungsziele untersucht. Ziel ist es, wiederkehrende Themen und Trends aus verschiedenen Freiwilligeneinsätzen zu identifizieren. Außerdem wird hinterfragt, welche Wahrnehmung und welches Interesse verschiedene Akteure (Partnerorganisation, Regierung, NRO etc.) an Freiwilligeneinsätzen verschiedener Art haben, um strategische Möglichkeiten und Partnerschaften für die Zukunft aufzuzeigen.

- Organisationsebene:

Bei diesem Workshop liegt der Schwerpunkt auf einem Erfahrungsaustausch zwischen verschiedenen Organisationen, die Freiwilligeneinsätze durchführen. Ziel ist es, Gemeinsamkeiten und Unterschiede in den Wirkungen von Freiwilligeneinsätzen zu diskutieren, um komparative Vorteile einer Maßnahme mit Freiwilligen gegenüber einer Maßnahme ohne Freiwillige zu identifizieren.

ABBILDUNG 13: Workshopebenen und -aktivitäten bei MCVD



b) Akteure

Die Workshops finden unter der Beteiligung folgender Akteure statt:

- Auf der Projektebene:
Freiwillige, lokale Projektbeteiligte und Partner inklusive Zielgruppe;
- Auf der Länderebene:
Freiwillige, regionale bzw. nationale Projektbeteiligte und Partner, evtl. Zielgruppe;
- Auf der Organisationsebene:
Organisationen, die Freiwilligeneinsätze durchführen, und Partner.

Die Wirkungsanalyse wird von Projektbeteiligten und Zielgruppenvertretern durchgeführt. Eine Validierung durch externe Experten sollte nach Möglichkeit Teil der Workshops sein. Diese externen Experten spielen während der Workshops die Rolle des „stillen Beobachters“ und geben gegen Ende

ihre eigene Reflektion zu den erarbeiteten Ergebnissen wieder. Der gesamte Prozess sollte von einem professionellen Moderator/Faszilitator begleitet werden.

c) Datenerhebung

Die Datenquellen sind Fragebögen und die Most Significant Change Stories, die von Freiwilligen vor den Workshops erarbeitet werden. Während des Workshops werden diese Ergebnisse von individuellen Freiwilligen vorgestellt und in der Gruppe kritisch hinterfragt und somit validiert. Je nach Workshopkontext werden dazu thematische oder regionale Fokusgruppen gebildet. Die Teilnehmer dieser Fokusgruppen sind Projektbeteiligte und Zielgruppenvertreter.

Auf einer solchen ersten Runde von Ergebnissen baut die darauf folgende Analyse auf. Sie strukturiert sich anhand folgender Fragestellungen:

- Welche der Initiativen waren am erfolgreichsten im Hinblick auf ihren Beitrag zum Entwicklungsziel A/ B/ C, and warum?
- Was war der spezifische Beitrag von Freiwilligen zum Entwicklungsziel A/ B/ C?
- Was waren die Hauptschwierigkeiten, warum?
- Welche waren die wichtigsten *Lessons Learned*?
- Auf welchem Level der Wirkungskette können die Beiträge von Freiwilligen eingeordnet werden?

Falls vorhanden, können während oder nach den Workshops Monitoringdaten zur Triangulation hinzugezogen werden.

d) Datenauswertung

Validation und Triangulation sind integraler Bestandteil eines jeden Workshops. Sie finden durch die gemeinsame Aufarbeitung und Analyse der Daten durch Projektbeteiligte und Zielgruppenvertreter statt. Konkret bedeutet dies, dass die oben genannten Fragen in Arbeitsgruppen bearbeitet und die Antworten geclustert und aggregiert werden. Die Formate für die Gruppenpräsentationen sind dabei festgelegt. In den meisten Fällen werden die Ergebnisse tabellarisch dargestellt.

Die synthetisierten Ergebnisse werden abschließend wiederum im Plenum vorgestellt. Ziel dabei ist es, durch die einzelnen Gruppenpräsentationen unterschiedliche Perspektiven und Wahrnehmungen gegenüber Freiwilligeneinsätzen herauszuarbeiten, um strategische Ansatzpunkte für die Zu-

kunft aufzudecken. Die Plenumsdiskussion garantiert, dass die Ergebnisse von den Workshopteilnehmern später anerkannt werden.

5. Nutzung der Ergebnisse

Da beim Workshop auf Länderebene die Beiträge von Freiwilligeneinsätzen weiter aggregiert werden, ist die Qualität der Dokumentation vorangegangener Workshops auf Projektebene entscheidend. Bei unzureichender Dokumentation sind die Ergebnisse für andere eventuell nur eingeschränkt nachvollziehbar. Für Workshopteilnehmer hingegen sind die Ergebnisse aller Erfahrung nach direkt verständlich.

UNV verwendete die Workshopergebnisse aus dem Jahr 2007 u. a. für die Berichterstattung gegenüber dem Verwaltungsrat. Da sich der zweite Teil der Workshops auf die Identifikation strategischer Ansatzpunkte für die zukünftige Arbeit konzentriert, liefern die Ergebnisse außerdem einen wichtigen Input zur programmatischen Arbeit der Organisation.

6. Anforderungen

Ein Workshop dauert typischerweise 3 Tage. Hinzu kommt die Zeit für Vorbereitung, u. a. die Beantwortung der Fragebögen durch Freiwillige. Außerdem nimmt die Nachbereitung Zeit in Anspruch, vor allem die sorgfältige Aufarbeitung der Workshopergebnisse (3-5 Tage).

Die Erfahrung hat gezeigt, dass ein guter Moderator/Faszilitator von zentraler Bedeutung für den Workshoperfolg ist. Nach entsprechendem Training kann das Verfahren auch von Projektverantwortlichen angewendet werden.

7. Besondere Eignung und Grenzen des Verfahrens

Die partizipativen Workshops bieten die Möglichkeit zu einer systematischen Reflektion über Wirkungen und den *added value* von Freiwilligeneinsätzen in der Entwicklungszusammenarbeit. Das gemeinsame Lernen mit Projektbeteiligten und Partnern steht dabei im Vordergrund. Insbesondere für die Freiwilligen bietet sich durch die Workshops die Gelegenheit, ihren eigenen Beitrag im größeren Kontext bzw. Entwicklungszusammenhang zu erkennen. Ihre Leistung wird anerkannt und honoriert, und die Ergebnisse aus den Workshops können einen wirksamen Beitrag zur Öffentlichkeitsarbeit im Bereich Freiwilligenengagement leisten.

Eine weitere Stärke des Verfahrens liegt in der Flexibilität und Vielseitigkeit der Anwendung. Je nach Kontext, Bedarf und Fragestellung ist es möglich, einzelne Workshopmodule getrennt zu nutzen.

Potenzielle Herausforderungen sind zum einen das unterschiedliche Verständnis von Freiwilligenarbeit in verschiedenen Kulturen und Ländern. Freiwilligeneinsätze bzw. freiwilliges Engagement kann viele unterschiedliche Facetten haben, wie z. B. Selbsthilfe auf lokaler Ebene, ehrenamtliche Tätigkeiten, zivilgesellschaftliches Engagement, Bürgerarbeit etc. Wie bei vielen qualitativen Ansätzen sind die Aggregation der Ergebnisse sowie die klare Zuordnung der Wirkungen zum Freiwilligeneinsatz eine potenzielle Schwäche des Verfahrens. Zum anderen kann es schwierig sein, manche qualitative Aspekte von Freiwilligeneinsätzen zu spezifizieren, wie z. B. Motivation und Engagement, Flexibilität, oder die Kenntnis lokaler Traditionen. Es ist daher wichtig, dass Teilnehmer zu Beginn der Workshops über gemeinsame Arbeitsdefinitionen für solche Begriffe und Konzepte verfügen.

Literaturhinweis

United Nations Volunteers (Hrsg.) (2008): A Participatory Methodology for Assessing the Contribution of Volunteering to Development. A Handbook for Volunteers and Programme Officers. Bonn.

Peace and Conflict Assessment (PCA)

MARTIN QUACK / DIRK SPRENGER

1. Allgemeine Informationen

Peace and Conflict Assessment (PCA) wurde in erster Linie von Thania Paffenholz und Luc Reyhler 2007 zur Analyse von Wirkungen auf Konflikte entwickelt. Andere Versionen des Verfahrens wurden unter den Namen Peace and Conflict Impact Assessment (PCIA) und Aid for Peace (AfP) veröffentlicht. PCA ist ein ausführlich (konzeptionell und methodisch) ausgearbeiteter Ansatz, um die im Übersektoralen Konzept „Krisenprävention, Konfliktbearbeitung und Friedensförderung in der deutschen Entwicklungszusammenarbeit“ des BMZ aufgezeigten Leitlinien auf der Programm- und Projektebene umzusetzen. Es steht damit für die deutsche EZ im Vordergrund und wird deshalb in diesem Steckbrief als Referenzverfahren vorgestellt. Elemente des Verfahrens wurden für die OECD DAC Draft Guidance for Peacebuilding Evaluations übernommen.

PCA wurde von der GTZ (u. a. PCIA im Südkaukasus 2003), der KfW, vom Life and Peace Institute (Schweden) für eine internationale *joint evaluation* in Mindanao und von der Friedrich-Ebert-Stiftung (u. a. in Afghanistan, Pakistan, Makedonien und Bosnien-Herzegowina) angewendet.

2. Anwendungsbereiche

Der Anwendungsbereich des Verfahrens umfasst sektoral solche Interventionen, die auf Konflikte einwirken sollen (*working on conflict*) oder das Potenzial dafür haben (*working in conflict*). Durch seine systematische Herangehensweise erleichtert PCA die konflikt sensible Ausrichtung der Projektarbeit einer Organisation und ist dabei nicht auf bestimmte Typen von Interventionen beschränkt. Aufgrund seines Umfangs ist es jedoch für kleinere NRO-Maßnahmen weniger geeignet. Im Vordergrund stehen Wirkungen auf den gesellschaftlichen Meso- und Makroebenen.

PCA ist ex-post möglich, aber auch zu Beginn, im Verlauf und am Ende einer Maßnahme. Das Verfahren ist in sich geschlossen und sollte idealerweise in den Projektzyklus integriert werden (eine der expliziten An-

wendungsformen des Verfahrens). PCA analysiert sowohl Prozesse als auch Zustandsveränderungen (abhängig vom vorhandenen Monitoringsystem).

3. Analyseansatz

PCA untersucht Wirkungen, die Einfluss auf Konflikt- und Friedensprozesse nehmen wollen. Es hat den Anspruch, einen umfassenden Rahmen zur Planung und Evaluation von Interventionen in Konfliktgebieten zu liefern. Die Analyse von Wirkungen von Interventionen zur Friedensförderung ist einer der Anwendungsbereiche des Verfahrens.

Prinzipiell sollen alle relevanten Wirkungen erfasst werden: geplante und ungeplante, direkte und indirekte. Außerdem wird untersucht, ob eine Intervention „relevant“ ist, d. h. ob ihr Ansatz zur Bedarfsanalyse passt.

Ansatzpunkt für die Wirkungsanalyse sind die Veränderungen im Wirkungsbereich der Intervention. Um diese Veränderungen feststellen zu können, sind *baseline*-Daten hilfreich. Anhand von Monitoringdaten zu Wirkungshypothesen mit Indikatoren wird abgeschätzt, welchen Einfluss die Intervention auf die festgestellten Veränderungen hatte.

Das Zuordnungsproblem wird nicht explizit erfasst, konkurrierende Ursachen für Veränderungen nicht explizit ausgeschlossen. Das Verfahren setzt sowohl auf (mehr oder weniger) objektive Monitoringdaten als auch auf subjektive Einschätzungen derjenigen Akteure, die an dem Verfahren beteiligt sind (partizipative Wirkungsbeobachtung). Kontrollgruppen sind nicht vorgesehen.

4. Durchführung

a) Verfahrensablauf

1. Teil: Friedens- und Konfliktanalyse, Analyse der Bedarfe für Friedensentwicklung

Grundlage des Verfahrens ist eine Analyse der Umwelt der Maßnahme (hinsichtlich der Bereiche Konflikt und Frieden), eine Trend- und Szenarioanalyse sowie eine Analyse der *peacebuilding vision* der Stakeholder und der Defizite der bestehenden Friedensförderung. Der konkrete Bedarf an *peacebuilding* soll auf dieser Grundlage spezifiziert werden.

2. Teil: Friedens- und konfliktbezogene Relevanzbewertung

Die Relevanz der Maßnahme wird durch den Vergleich ihrer Ziele und Aktivitäten mit dem identifizierten Bedarf bewertet, z. B. mit Hilfe einer Relevanz-Skala. Dabei sollen ähnliche Maßnahmen anderer Akteure und *lessons learned* beachtet werden.

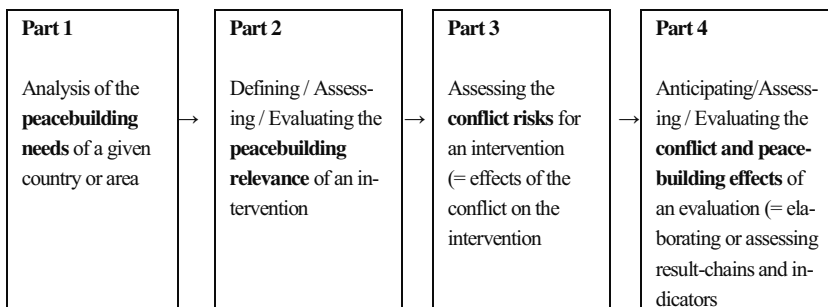
3. Teil: Risikomanagement

Die Risiken, denen eine Maßnahme durch den Konflikt ausgesetzt ist, werden mit Hilfe von Checklisten zu Sicherheit, politischem Klima, Verhältnis zu Partnern und Stakeholdern sowie zu Konfliktparteien und anderen intervenierenden Akteuren eingeschätzt. Eine Aufteilung nach Regionen ist möglich.

4. Teil: Friedens- und konfliktbezogene Wirkungsbeobachtung

Mit Hilfe von Wirkungshypothesen, *baseline*-Daten (die ggf. rekonstruiert werden sollen) und Monitoring-Daten wird die aktuelle Situation mit früheren Zeitpunkten verglichen, um Veränderungen zu identifizieren. Außerdem werden verschiedene Methoden zur spezifischen Wirkungsanalyse angeboten (u. a. Kriterien von Do No Harm und von Reflecting on Peace Practice (www.cdainc.com) für Wirkungen auf Makroebene). Empfohlen werden *joint evaluations*, um einen Teil der konkurrierenden Ursachen auszuschließen.

ABBILDUNG 14: Aid for Peace Framework



Quelle: Paffenholz/Reychler (2007), S. 17

b) Akteure

Welche Akteure wie an dem Verfahren beteiligt sind ist nicht festgelegt und hängt von der Art der Anwendung ab. Die Beteiligung von Stakeholdern in verschiedenen Phasen wird empfohlen. Aufgrund des komplexen Vorgehens ist es sinnvoll, das Verfahren entweder in den Projektzyklus zu integrieren oder mit externer Unterstützung durchzuführen. Prinzipiell sind auch Selbstevaluationen möglich. Die Integration in den Projektzyklus ist auch sinnvoll für *capacity building*.

c) Datenerhebung

Die Datenquellen für PCA sind flexibel zu kombinieren. In erster Linie handelt es sich um *baseline*- und Monitoringdaten sowie Interviews. Ein *sampling* ist nicht vorgesehen. Die Instrumente zur Datenerhebung hängen von der vermuteten Wirkungsweise und der Wirkungsebene (z. B. Stärkung bestimmter gesellschaftlicher Gruppen) ab. Denkbar sind u. a. Workshops, Interviews und Umfragen. Unterschiedliche Perspektiven und Wahrnehmungen werden im Datenerhebungsverfahren nicht explizit berücksichtigt.

d) Datenauswertung

Die Daten werden in erster Linie durch Workshops ausgewertet (je nach Art der Durchführung). Das Auswertungsraster wird in den ersten Teilen des Vorgehens erarbeitet. Unterschiedliche Perspektiven und Wahrnehmungen werden implizit in den Workshops berücksichtigt. Methoden zur Validierung des Ergebnisses sind in PCA nicht festgelegt.

5. Nutzung der Ergebnisse

Die logische Abfolge der einzelnen Teile fördert die Transparenz, Verständlichkeit und Nachvollziehbarkeit der Ergebnisse. Eine Quantifizierung oder Aggregation der Ergebnisse ist nicht vorgesehen (aber auch nicht ausgeschlossen). Zur optimalen Nutzung der Ergebnisse wird die Integration von PCA in den Projektzyklus empfohlen. Die Beteiligten sollen PCA vor allem zur Steuerung der Maßnahme nutzen.

6. Anforderungen

Da eine Einführung in das Verfahren notwendig ist, beträgt der Zeitaufwand für PCA bei erstmaliger Anwendung mindestens 2 Tage (abhängig von ver-

fügbarer Vorinformation), im Rahmen des Monitoring ggf. weniger. Je nach Entwicklung des Konflikts sollte PCA wiederholt durchgeführt werden.

Eine Moderation ist notwendig, je nach Fokus intern oder extern. Ferner können, je nach Programmumfang, externe Gutachter/-innen zum Einsatz kommen. Darüber hinaus entstehen keine Kosten.

Baseline- und Monitoringdaten sowie formulierte Wirkungsketten sind sehr hilfreich für die Durchführung von PCA.

7. Besondere Eignung und Grenzen des Verfahrens

Zu den Stärken von PCA zählt die systematische Herangehensweise, die Möglichkeit zur Integration verschiedener Elemente (z. B. DNH), und dass das Verfahren für viele verschiedene Anwendungen angepasst werden kann. Es dient außerdem der Konfliktsensibilisierung der Beteiligten.

Als Schwächen von PCA lassen sich die vielen Auswahlmöglichkeiten anführen. Speziell zur Wirkungsanalyse gibt es kaum klare Vorgaben. Es gibt eine Tendenz, zu viele Beobachtungsfelder und -ebenen in den Fokus zu nehmen und das Verfahren damit zu überfrachten. Es handelt sich um eine relativ zeitaufwändige Methode, die ihre volle Wirkung erst bei regelmäßiger Anwendung im Rahmen des Projektzyklus entfaltet. PCA tendiert außerdem zu einer Linearität (und geringen Reflexivität) in der Durchführung: Beispielsweise wird die Relevanz einer Maßnahme schon im zweiten Schritt bestimmt, obwohl erst der vierte Schritt des Verfahrens dazu wertvolle Erkenntnisse liefert.

Literaturhinweise

- Bush, Kenneth (1998): A Measure of Peace: Peace and Conflict Impact Assessment of Development Projects in Conflict Zones (Working Paper No.1). The Peacebuilding and Reconstruction Programme Initiative and The Evaluation Unit, IDRC.
- GTZ (2007): Peace and Conflict Assessment (PCA): Ein methodischer Rahmen zur konflikt- und friedensbezogenen Ausrichtung von EZ-Maßnahmen. Eschborn. (www.2.gtz.de/dokumente/bib/07-1526.pdf).
- Paffenholz, Thania und Reychler, Luc (2007): Aid for Peace: A Guide to Planning and Evaluation for Conflict Zones. Baden-Baden: Nomos.
- Paffenholz, Thania (2005): Third-generation PCIA : Introducing the Aid for Peace Approach. (www.berghof-handbook.net/uploads/download/dialogue4_paffenholz.pdf).
- Wouters, Shirley et al. (2003): Development of a Peace and Conflict Impact Assessment for Communities in the South Caucasus. Berlin: SLE (Schriftenreihe des Seminars für ländliche Entwicklung Nr. S206).

- Berghof Research Center for Constructive Conflict Management (2005): Berghof Handbook for Conflict Transformation, Dialogue Series No. 4 New Trends in PCIA. Berlin. (www.berghof-handbook.net/uploads/download/dialogue4_pcia_new_complete.pdf).
- Friedrich-Ebert-Stiftung (2007): Peace and Conflict Impact Assessment Methodical Guidelines. Bonn. (<http://library.fes.de/pdf-files/iez/05594-info.html>).

Programm-integriertes Planungs-, Monitoring- und Evaluierungssystem (PriME) von InWent

KAROLA BLOCK / SONJA NELLES

1. Allgemeine Informationen

PriME – Programm-integriertes Planungs-, Monitorings- und Evaluierungs-(PM&E-)System von InWent – wurde entwickelt, um den internationalen Herausforderungen an Wirkungsorientierung und -erfassung im Rahmen von Capacity Building Programmen (Personalentwicklungsberatung, Training, Dialog und Netzworkebildung) gerecht zu werden. Das PM&E-System berücksichtigt die Anforderungen an Wirkungsorientierung von Programmen im Rahmen der internationalen Personal- und Organisationsentwicklung und orientiert sich an nationalen und internationalen Standards wie z. B. den Standards und Kriterien des DAC der OECD und den Standards für Evaluation der DeGeval.

PriME wurde in den Jahren 2004 und 2005 unter Einbindung interner und externer Expertise entwickelt. Das System wird kontinuierlich auf der Basis praktischer Anwendungen weiterentwickelt. PriME ist das verbindliche PM&E-System für die BMZ-Geschäftsbereiche von InWent. In anderen Geschäftsbereichen für andere Auftraggeber laufen sukzessive und an die spezifischen Rahmenbedingungen angepasste Einführungen.

PriME leistet einen wichtigen Beitrag zur Qualitätssicherung der Programme von InWent und stellt eine wesentliche Komponente des Qualitätsmanagements sowie der Programmsteuerung von InWent dar. Das System fördert eine Qualitätsverbesserung bei gleichzeitiger Arbeitserleichterung durch Standardisierung und Systematisierung von Planungs-, Monitoring- und Evaluierungsaktivitäten und leitet kontinuierliche Verbesserungsprozesse durch institutionelles Lernen ein. Damit ist PriME auch ein Instrument zur Schaffung von Transparenz und zur Stärkung der Legitimation gegenüber Auftraggebern und Öffentlichkeit.

Es wurde ein PriME-Handbuch entwickelt. Darüber hinaus ist das Handbuch mit *best* und *good practices* zu den einzelnen PriME-Instrumenten im Rahmen von PM&E im InWent-Intranet abgebildet.

Verantwortlich für die (Weiter-)Entwicklung und Einführung ist die Abteilung 1.04 Qualität und Evaluierung bei InWent. Externe Beratung er-

folgt bei der Konzeption, dem Umsetzungsprozess und verschiedenen internen Trainingsmaßnahmen. Anwender/-innen sind in erster Linie InWEnt-Mitarbeiter/-innen. Darüber hinaus werden punktuell externe Fachpartner wie z. B. Gutachter/-innen eingebunden, z. B. im Rahmen von Evaluierungen.

2. Anwendungsbereiche

Im Mittelpunkt von PriME steht die Erfassung und der Nachweis der erreichten Ziele und Wirkungen von InWEnt-Programmen in allen Sektoren und auf allen Ebenen (Mikro-, Meso- und Makroebene). Das PM&E-System berücksichtigt dabei jeweils die im InWEnt-Capacity-Building-Ansatz definierten Interventionsebenen Individuum, Organisation und System. Im Fokus steht die Förderung von individuellen Kompetenzen von Führungs- und Fachkräften aus Wirtschaft, Staat und Zivilgesellschaft, die für den Veränderungsbedarf von Organisationen und Gesellschaften benötigt werden.

PriME deckt alle Phasen in einem Projektzyklus ab: Planung, Durchführung im Sinne von Monitoring und Evaluierung. PriME umfasst dabei die Erfassung und Bewertung von Zuständen und Prozessen (Veränderungsprozessen) u. a. durch ein projektbegleitendes Monitoring.

3. Analyseansatz

Im Rahmen von PriME werden v. a. geplante und mit Indikatoren versehene direkte, kurz- und mittelfristige Wirkungen auf *outcome*-Ebene erfasst und bewertet. Weiterhin wird die Qualität der einzelnen Maßnahmen (z. B. Training, Dialogveranstaltungen) systematisch gemessen. Indirekte, längerfristige Wirkungen auf der *impact*-Ebene werden über die Darstellung von Plausibilitäten und Wirkungszusammenhängen abgeleitet und eingeschätzt.

Ansatzpunkte für die Erfassung und Bewertung von Wirkungen im Rahmen von PriME sind die Interventionen, d. h. die konkreten Maßnahmen. Ziel ist festzustellen, in welchem Maße Veränderungen stattfinden, die diesen konkreten Maßnahmen zugeordnet werden können. Da der InWEnt-spezifische Beitrag im Rahmen von Capacity Building der Aufbau und die Förderung von individuellen Handlungs- und Veränderungskompetenzen ist, liegt der Schwerpunkt bei Monitoring und Evaluierung von Wirkungen in der Überprüfung, inwieweit die Teilnehmer/-innen an den InWEnt-Maßnahmen das Gelernte im Arbeitskontext im Sinne der angestrebten Veränderungen auf Institutionen- und Systemebene anwenden.

Zur Umsetzung der Wirkungsorientierung setzt PriME auf der Ebene der direkten Wirkungen / *outcome* an und geht von einer direkten Ursache-Wirkungs-Beziehung aus. Daher wird schon in der Planung der Wirkungszusammenhang zwischen Maßnahmen bzw. deren Outputs und dem jeweiligen *outcome* (entsprechend GTZ-Terminologie: Nutzung der Leistung und direkter Nutzen) in den Zielformulierungen definiert. Z. B. „Führungskräfte von Kommunen haben Kompetenzen zu Verfahren der Bürgerbeteiligung erworben (Output) und leiten in der Folge partizipative Planungsverfahren in ihren Kommunen ein (*outcome*)“.

Veränderungen, die dadurch auf der Systemebene ausgelöst werden, werden der Impact-Ebene zugeordnet, eine eindeutige und direkte Zuordnung dieser längerfristigen Wirkungen zu den InWEnt-Maßnahmen ist nicht mehr möglich. In der Praxis sind alle Veränderungen nicht immer eindeutig nur einer InWEnt-Maßnahme zuordenbar, sondern auch Folge unterschiedlicher Interventionen, verstärkt durch die Bemühungen einer „EZ aus einem Guss“. Methodisch kann bei PriME sowohl mit „faktischen“ Wirkungsnachweisen als auch mit subjektiven Einschätzungen von Projektbeteiligten gearbeitet werden.

4. Durchführung

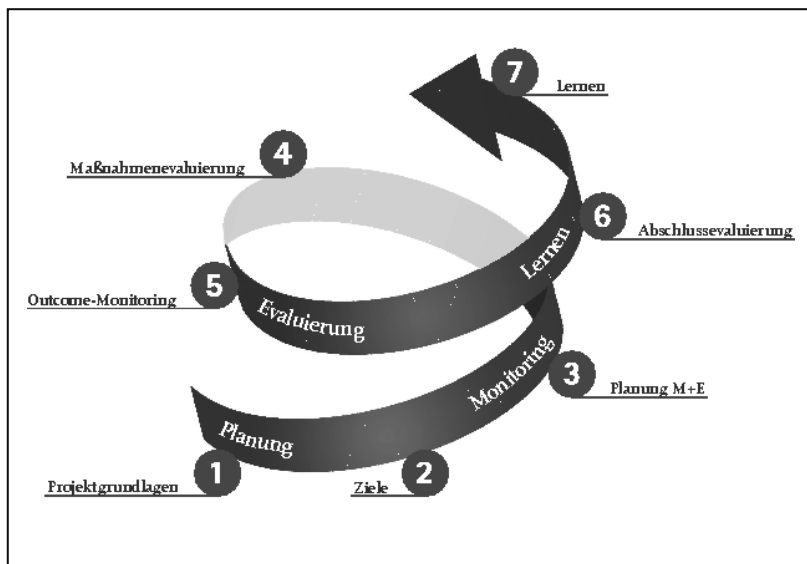
a) Verfahrensablauf

Der einheitliche PM&E-Zyklus mit den einzelnen Arbeitsschritten im Rahmen von PriME ist nachfolgend dargestellt. Dieser Zyklus bezieht sich auf die gesamte Programmlaufzeit und stellt keinen jährlichen Programmzyklus dar.

b) Akteure

Verantwortlich für die Anwendung und Umsetzung von PriME sind die Programmabteilungen, insbesondere die InWEnt-Projektleitungen. Die Einbindung der Projektpartner vor Ort in alle Schritte im Rahmen von PM&E ist integraler Bestandteil aller InWEnt-Programme. Der Grad der Einbindung ist unterschiedlich, z. T. wird die Datenerhebung, Analyse und Dokumentation im Rahmen von M&E ganz in die Verantwortung der Partner gegeben und entsprechend formal über ein *memorandum of understanding* abgesichert. Allerdings ist es häufig noch schwierig, wirkungsorientierte Informationen (und nicht nur Berichte über durchgeführte Leistungen) von Partnerorganisationen zu erhalten.

ABBILDUNG 15: PM&E-Zyklus und Arbeitsschritte bei PriME



Im Rahmen von PriME erfolgt das Outcome-Monitoring i. d. R. durch die Projektleitungen und die beteiligten Partner, der Einsatz von Gutachtern ist optional. Abschlussevaluierungen und Ex-Post-Evaluierungen werden dagegen standardmäßig extern durchgeführt.

c) Datenerhebung

Für die Maßnahmenevaluierung als Qualitätssicherungsinstrument im Rahmen von PriME werden teil-standardisierte Fragebögen für verschiedene Instrumente (z. B. E-learning, Training, Dialog) eingesetzt. Diese bestehen aus 11 verpflichtenden Fragen sowie einem Set an Empfehlungen für die weitere Fragebogengestaltung. In der Regel richten sie sich an Teilnehmende zur direkten Bewertung von Trainings- oder Qualifizierungsveranstaltungen und deren Nutzen für Veränderungen im Arbeitsbereich.

Für das Monitoring des Outcome sowie für Abschlussevaluierungen gibt PriME die Methoden für die Datenerhebung nicht vor, da Wirkungen auf der relevanten Outcome-Ebene nur projektspezifisch formuliert werden können. Es werden jedoch geeignete Methoden wie z. B. online-basierte Umfragen, Interviews, Planungs- oder Monitoringworkshops usw. im Pri-

ME-Handbuch nach einem einheitlichen Raster beschrieben und im Hinblick auf ihre Eignung bewertet. Konkrete projektspezifische Beispiele werden laufend im Intranet veröffentlicht.

d) Datenauswertung

Die Daten der Maßnahmenevaluierung werden entweder mit dem Programm „Grafstat“ oder anderen quantitativen oder teil-quantitativen Verfahren ausgewertet. Die Ergebnisse fließen in das Monitoring des Outcome mit ein und werden in diesem Prozess mit anderen Methoden trianguliert, d. h. nochmals kontrolliert. Unabhängig von den im Einzelfall verwendeten Methoden der Informationsbeschaffung besteht beim Outcome-Monitoring der Anspruch, die unterschiedlichen Perspektiven und Interessen der Akteure zu erfassen und ihre Relevanz bezüglich des weiteren Verlaufs der Vorhaben zu analysieren.

5. Nutzung der Ergebnisse

Die Nutzung von Instrumenten und Ergebnissen von PriME sind in allen folgenden Schritten im Rahmen des Projektmanagement angelegt:

- Planungsinstrumente und Standardformate dienen der Konzeption und Beantragung von Projekten.
- Maßnahmenevaluierungen werden für die unternehmensweite und projektspezifische Qualitätssicherung genutzt.
- Regelmäßiges Outcome-Monitoring dient der Projektsteuerung im Rahmen der jährlichen Operationsplanung und zur Erfüllung der Berichtspflichten gegenüber den Auftraggebern.
- Die Abschlussevaluierungen dienen dem institutionellen Lernen als Auftrag an Projekt- und Abteilungsleiter/-innen und der Erfüllung der Berichtspflicht gegenüber den Auftraggebern. Im Rahmen der Abschlussevaluierung wird eine Gesamtbewertung der Wirkungen der Programme vorgenommen sowie *lessons learned* für die Gestaltung weiterer Programme formuliert. Ein wichtiger Bestandteil der Abschlussevaluierung ist eine Analyse der Einflussfaktoren, die in erster Linie als Lern- und Wissensinstrument dient. Sie hilft, erfolgreiche und kritische Aspekte herauszufinden, um die Planung von weiteren Vorhaben gezielter vornehmen zu können.
- Abschlussevaluierungen werden darüber hinaus bei übergeordneten Wirkungsanalysen und ex-post-Evaluierungen zum Nachweis von aggregierten InWEnt-Wirkungen auf Länder- bzw. Sektorebene be-

rücksichtigt. Wirkungsanalysen und ex-post-Analysen sind Bestandteil eines unabhängigen zweijährigen Evaluierungsprogramms, welches von der Abteilung 1.04 Qualität und Evaluierung verantwortet und durchgeführt wird.

Um eine Vergleichbarkeit aller Evaluierungsformate sicherzustellen, gibt es ein einheitliches Berichtsraster. Für Wirkungsanalysen und ex-post-Analysen wird zukünftig ein Rating-System verwendet, welches derzeit noch entwickelt wird.

6. Anforderungen

Obwohl PriME Standards und Verfahrensschritte vorgibt, ist die Anwendung von Methoden bzw. die Ausgestaltung einzelner Arbeitsschritte flexibel und an die Rahmenbedingungen und Anforderungen der Programme anpassbar. Da InWent vor Ort nur in wenigen Ländern über eine institutionelle Basis verfügt, stellt insbesondere das Outcome-Monitoring InWent trotzdem vor erhebliche Herausforderungen.

Die Kosten für PM&E werden im Projektbudget von Anfang an budgetiert; international werden dafür durchschnittlich 5-10% angesetzt.

Die Anwendung von PriME wird kontinuierlich im Rahmen von sogenannten Werkstatt-Tagen vermittelt und die Anwendung der Instrumente am jeweiligen Projekt geübt. Parallel dazu gibt es individuelle Beratungsmöglichkeiten.

7. Besondere Eignung und Grenzen des Verfahrens

Besondere Stärken von PriME sind auf der einen Seite, dass hier Mindestqualitätsanforderungen für das PM&E von Programmen gestellt werden und diese gleichzeitig ausreichend Spielraum für die projektspezifische Ausgestaltung bieten. PriME berücksichtigt die spezifischen Anforderungen an Personal- und Organisationsentwicklung im Rahmen von *capacity building*, um Wirkungen auf verschiedenen Interventionsebenen (Capacity-Building-Ansatz von InWent: Individuum – Institutionen – Systemebene) zu erzielen. Dabei verfolgt es eine konsequente Differenzierung der Wirkungsebenen nach der DAC-Terminologie.

PriME beinhaltet die Chance zu stärkerer Partnerorientierung (gerade als EZ-Institution mit zentralisierter Struktur) und zu mehr Transparenz und gemeinsamem Lernen. Daher kann PriME für andere Institutionen durchaus hilfreiche Orientierung und Instrumente bieten. Inwieweit diese Stärke in

der Praxis zum Tragen kommt, muss sich erst noch erweisen, da PriME erst seit 2005 systematisch im BMZ-Geschäftsbereich angewendet wird.

Literaturhinweise

InWEnt (2006): PriME – Programm-integriertes Planungs-, Monitoring- und Evaluierungssystem von InWEnt. Handbuch, Abt. 1.04 Qualität und Evaluierung. Bonn [überarbeitete und ergänzte Version liegt seit Dezember 2008 vor und ist auch im Internet verfügbar].

Quantitative experimentelle Designs

STEFAN KLONNER

1. Allgemeine Informationen

Experimentelle Methoden gehören zur Gruppe der quantitativen Verfahren der Wirkungsanalyse. Im Vordergrund steht die Bestimmung eines Wirkungseffekts (Treatment Effect) einer bestimmten Maßnahme. Von primärem Interesse ist hierbei der individuelle Wirkungseffekt, der als der Unterschied zwischen zwei Szenarien definiert ist, der Situation eines Individuums nach Teilnahme an einer bestimmten Maßnahme und der Situation des selben Individuums, wenn es an der Maßnahme nicht teilgenommen hätte. Der durchschnittliche Wirkungseffekt ist hingegen der Durchschnitt der individuellen Wirkungseffekte in einer gegebenen Population von Individuen. Die grundlegende Eigenschaft von experimentellen Designs ist, dass Teilnehmer (Treatment Group) und Nicht-Teilnehmer (Control Group) an einer Maßnahme zufällig vor ihrer Durchführung bestimmt werden und die Output- und Outcome-Variablen von Interesse in beiden Gruppen vor und nach Durchführung der Maßnahme erhoben werden. Für eine gegebene Situation lässt sich damit der durchschnittliche Wirkungseffekt ermitteln, der den entsprechenden Output bzw. Outcome der Maßnahme quantifiziert. Die methodische Bedeutung experimenteller Designs liegt darin begründet, dass im rigoros wissenschaftlichen Denken nur sie eine konsistente Schätzung durchschnittlicher Wirkungseffekte erlauben, denn nur bei ihnen wird systematisch mit Kontrollgruppen und mit großen Fallzahlen gearbeitet. Wenn bei der Projektplanung kein experimenteller Ansatz zum Einsatz kommt, unterscheiden sich Treatment- und Control-Gruppe im Allgemeinen von vornherein in einer Weise, die eine gültige Schätzung eines Wirkungseffektes unmöglich macht. Entsprechend kommen alle anderen quantitativen, nicht-experimentellen Designs, die der Schätzung von Wirkungseffekten dienen, nicht ohne zusätzliche Annahmen aus. Experimentelle Methoden haben im Verlauf der letzten Jahre auch im Rahmen der Evaluierung entwicklungspolitischer Maßnahmen an Bedeutung gewonnen. Bei Projekten der Weltbank, verschiedener nationaler Programme und zahlreicher Nicht-Regierungs-Organisationen kommen sie heute häufiger zur Anwendung. Darüber hinaus werden experimentelle Designs zunehmend von akademi-

schen Ökonomen, auch führender US-amerikanischer Universitäten (Harvard, Massachusetts Institute of Technology, Yale, Berkeley), angewandt.

2. Anwendungsbereiche

Eine grundlegende Voraussetzung der Anwendung experimenteller Designs ist, dass zufällig ausgewählte Treatment- und Control-Individuen in ausreichender Anzahl zur Verfügung stehen und dass die Kontrollgruppe keinen, auch keinen indirekten, Effekten der Maßnahme unterliegt. Diese Anforderungen machen sofort klar, dass sich experimentelle Methoden nicht zur Evaluierung von Maßnahmen eignen, die auf der Makroebene angesiedelt sind. Hingegen sind Projekte auf der Mikroebene und z. T. auf der Mesoebene gut geeignet. In der Vergangenheit wurden Gesundheits- und Erziehungsmaßnahmen, Programme zur Förderung kleinster und kleiner Unternehmer, Entwicklung des ländlichen Raums, aber auch Reformen der Kommunalregierung erfolgreich evaluiert.

3. Analyseansatz

Im Vordergrund steht die quantitative Bestimmung durchschnittlicher Wirkungseffekte. Bei entsprechend sorgfältiger Konzeption und Anwendung des Evaluationsverfahrens lassen sich darüber hinaus auch hypothetische Wirkungsketten überprüfen. Das wichtigste Evaluationsergebnis besteht aus Tabellen, in denen für verschiedene Variablen die der Maßnahme zuzuschreibenden Veränderungen sowie deren Standardfehler aufgeführt sind. Dies erlaubt neben der Schätzung von durchschnittlichen Effekten die Beantwortung der Frage, ob eine Maßnahme einen statistisch signifikanten Effekt auf eine bestimmte Ausgangssituation hat. Diese einfache, quantitative Darstellung der Ergebnisse erlaubt einen Vergleich mit der Wirkung verwandter oder alternativer Maßnahmen und kann somit eine nützliche Grundlage für größer angelegte Kosten-/Nutzenanalysen bilden.

4. Durchführung (inkl. Datenerhebung und Datenauswertung)

a) Verfahrensablauf

Experimentelle Methoden bei der Evaluierung setzen eine experimentelle Anlage der Maßnahme bzw. des Projekts voraus. Dies bedeutet, dass die Maßnahme und ihre Evaluation simultan geplant und durchgeführt werden müssen. Wichtigster Punkt hierbei ist die zufällige Auswahl der Teilnehmer

und der Kontrollindividuen vor Beginn der Maßnahme. Es folgt ein schematischer Ablauf einer auf experimentellen Methoden basierenden Evaluation.

1. Projekt- und Evaluationsplanung

Vor Beginn der Maßnahme findet zweierlei statt. Zum einen werden Variablen identifiziert, auf die das geplante Projekt möglicherweise einen Einfluss hat. Die Auswahl der Variablen, deren Änderung gemessen werden soll, kann aus einer vorhergehenden Erstellung hypothetischer Wirkungsketten resultieren. Solche Überlegungen sind insbesondere zur Vermeidung von späteren Zuordnungslücken erforderlich. Die Variablen können quantitativer oder qualitativer Art sein. Zum anderen werden die Treatment- und die Control-Gruppe ausgewählt. Diese werden aus einer Grundgesamtheit von Beobachtungseinheiten zufällig bestimmt oder „randomisiert“. Wichtig ist, die Ebene der Beobachtungseinheiten, auf der randomisiert wird, so zu wählen, dass Control-Individuen nicht von der Maßnahme beeinflusst werden. Um dies zu erreichen, können teilweise einfache Wege gefunden werden. Nehmen wir das Beispiel eines Schulbücherprogramms, dann ist z.B. eine Randomisierung auf Schul-Ebene (d. h. 50% der Schulen erhalten Bücher, 50% keine) sinnvoller als auf individueller Ebene innerhalb der Schulen (in jeder Schule erhalten genau 50% der Schüler Bücher), denn während im zweiten Szenario die Kontroll-Individuen von den Büchern ihrer Mitschüler profitieren könnten, ist das im ersten Szenario sehr unwahrscheinlich. Bleiben solche Spillover-Effekte unberücksichtigt, könnte dies eine Unterschätzung des durchschnittlichen Wirkungseffekts zur Folge haben. Wenn die statistische Signifikanz der Ergebnisse von Bedeutung ist (z. B.: „Eine Verbesserung der Lesefähigkeit von 5% oder mehr soll mit 95-prozentiger Wahrscheinlichkeit als statistisch signifikant identifiziert werden), kann die erforderliche Stichprobengröße mit Hilfe von Güteberechnungen ermittelt werden. Hierzu sind Daten über die Verteilung (Mittelwert und Varianz) der betrachteten Variablen in der Population erforderlich (siehe Duflo et al., 2008, Abschnitt 4). Ist eine Disaggregation der Ergebnisse (z. B. nach Gender) geplant, muss dies auch bei der Stichprobengröße berücksichtigt werden.

2. Baseline Survey

Vor Durchführung der Maßnahme werden die Variablen von Interesse bei allen Individuen der Stichprobe erhoben. Die Randomisierung stellt im Prinzip sicher, dass sich die Treatment- und Control-Individuen nicht in sta-

tistisch signifikanter Weise unterscheiden. Dies sollte aufgrund der *baseline*-Daten überprüft werden.

3. Durchführung der Maßnahme

Eventuell begleitende Datenerhebungen in der Stichprobe, gemäß Schritt 1 (z. B. Erhebung der Nutzung der ausgeteilten Schulbücher bei den *treatment*-Schülern).

4. Follow-up Survey

Nach Durchführung der Maßnahme erneute Erhebung der Variablen von Interesse bei allen Individuen der Stichprobe.

5. Statistische Auswertung

Der durchschnittliche Wirkungseffekt auf Variable Y (z. B. Punktzahl bei einem Lesetest) wird berechnet als doppelte Differenz, $(\bar{y}_1^{\text{treat}} - \bar{y}_0^{\text{treat}}) - (\bar{y}_1^{\text{control}} - \bar{y}_0^{\text{control}})$, wobei $\bar{y}_t^{\text{group}}$ den Stichprobenmittelwert der Untergruppe $\text{group} = \{\text{treat}, \text{control}\}$ in Survey-Runde t (0 für *baseline*, 1 für *follow-up*) bezeichnet. Mit Hilfe elementarer Methoden der linearen Regressionsanalyse lassen sich darüber hinaus auf einfache Weise die statistischen Eigenschaften des durchschnittlichen Wirkungseffekts bestimmen.

b, c, d) Akteure, Datenerhebung und -auswertung

Die Datenerhebung kann vom Projektteam oder einer lokalen Survey-Organisation durchgeführt werden. Um die Objektivität der Ergebnisse zu gewährleisten, ist eine Trennung der Projektdurchführung einerseits und der Datenerhebung und -auswertung andererseits jedoch wünschenswert.

5. Nutzung der Ergebnisse

Experimentelle Methoden betonen die Konsistenz und Objektivität der gewonnenen statistischen Ergebnisse. Entsprechend bieten die Ergebnisse nicht nur der durchführenden Organisation oder dem Geber einen Überblick über die Wirkung der Maßnahme. Insbesondere kann das Wissen um die Wirkung bzw. den Erfolg bestimmter Maßnahmen als öffentliches Gut für die Gestaltung zukünftiger Maßnahmen angesehen werden. Entsprechend argumentieren Befürworter dieses Ansatzes, dass auf experimentellen Methoden beruhende Evaluationsergebnisse in einer Datenbank gesammelt und

als Grundlage künftiger Projektplanung dienen sollten. Ansätze hierzu sind bei der Weltbank vorhanden. Die Organisation International Initiative for Impact Evaluation (3ie) verfolgt explizit dieses Ziel und bietet eine Plattform, auf der entsprechende Ergebnisse online verfügbar gemacht werden.

6. Anforderungen

Die Anwendung experimenteller Methoden ist im Allgemeinen aufwändig. Sie erfordert eine randomisierte Durchführung der Maßnahme und eine genaue Vorabplanung.

Literaturhinweise

- Duflo, Esther; Glennerster, Rachel und Kremer, Michael (2008): Using Randomization in Development Economics Research: A Toolkit. Kapitel 61 in: Schultz, T. Paul und Strauss, John (Hrsg.): Handbook of Development Economics, vol. 4, pp 3895-3962. Elsevier: North-Holland.
- Duflo, Esther und Kremer, Michael (2005): Use of Randomization in the Evaluation of Development Effectiveness. In: Pitman, George; Feinstein, Osvaldo und Ingram, Gregory (Hrsg.): Evaluating Development Effectiveness. New Brunswick, NJ: Transaction Publishers 2005, pp. 205-232.
- Glewwe, Paul; Kremer, Michael und Moulin, Sylvie (2004): Many Children Left Behind? Textbooks and Test Scores in Kenya. American Journal: Applied Economics, 1(1): 112-135.
- Kremer, Michael (2003): Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. American Economic Review Papers and Proceedings 93(2): 102-115.
- Ravallion, Martin (2001): The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation. World Bank Economic Review 15(1), 115-14.
- Ravallion, Martin (2008): Evaluating Anti-Poverty Programs. Kapitel 59 in Schultz, T. Paul und Strauss, John (Hrsg.): Handbook of Development Economics, vol. 4, pp 3787-3846. Elsevier: North-Holland.

Quantitative nicht-experimentelle Designs

VON MARKUS FRÖLICH

1. Allgemeine Informationen

Ziel der quantitativen Evaluation ist die Abschätzung der Wirkung einer entwicklungspolitischen Maßnahme mit Hilfe von statistischen Verfahren, die nicht nur die Größenordnung der Effekte quantifizieren sondern auch die statistische Schätzungenauigkeit (Konfidenzintervall) berechnen. Neben der Berücksichtigung von Eigenschaften der Datenerhebung werden komplexere Verfahren insbesondere zur Handhabung des Selektionsproblems bei der Definition der Kontrollgruppen benötigt. Dieses Problem wird hier zunächst diskutiert. Alle beschriebenen Verfahren basieren auf einem Kontrollgruppenansatz. Dabei werden alle Ansätze, bei denen keine völlig zufällige Zuteilung erfolgt, als nicht-experimentelle Ansätze bezeichnet; gleichwohl sind einige der nachfolgenden Verfahren, insbesondere Instrumentalvariablenmethoden, in der Praxis nahezu immer auch bei experimentellen Ansätzen erforderlichlich.

Zur Beschreibung der quantitativen Evaluationsansätze ist es hilfreich, in beispielhafter Form die grundlegende Notation einzuführen. Wir betrachten die Wirkung einer Intervention D auf eine Ergebnisvariable Y . Die Intervention oder entwicklungspolitische Maßnahme ist z. B. die Bohrung eines Brunnens, die Bereitstellung von sauberem Wasser, eine Schutzimpfung, die Bereitstellung von Schulbüchern, eine bestimmte Entlohnungsform für Lehrer, eine dezentralisierte Gemeindeverwaltung etc., um nur einige Beispiele zu nennen. Ein Individuum i oder z. B. eine Gemeinde nimmt entweder an der Intervention teil oder nicht. Wir betrachten hier nur den binären Fall, wo es nur eine Intervention gibt, d. h. Teilnahme ja oder nein. Eine Erweiterung des Modellrahmens ist jedoch leicht möglich. Mit der Variable Y wird eine Zielgröße der Intervention gemessen, z. B. Gesundheitszustand, schulische Leistungen. Eine Erweiterung auf multiple Zielgrößen, intermediäre Ergebnisvariablen und die Analyse von Wirkungsketten ist ebenfalls leicht möglich.

Der Effekt der Intervention ist nun definiert als die Differenz zwischen dem hypothetischen Ergebnis Y_i^1 , das sich ergeben würde, wenn Individuum i an der Maßnahme teilnähme, und dem hypothetischen Ergebnis Y_i^0 ,

welches sich ergeben würde, wenn dieses Individuum i nicht an der Maßnahme teilnähme. Der individuelle Wirkungseffekt der Maßnahme ist definiert als $Y_i^1 - Y_i^0$, also das Ergebnis nach Teilnahme an der Maßnahme relativ zu dem Effekt, der sich für Individuum i im Falle einer Nichtteilnahme ergeben hätte. Die Wirkung einer Maßnahme wird oftmals im Durchschnitt über die Population geschätzt und ist definiert als der durchschnittliche Maßnahmeneffekt $E[Y^1 - Y^0]$ oder als der durchschnittliche Effekt für die Teilnehmer $E[Y^1 - Y^0 | D = 1]$.

Das fundamentale Problem ist nun, dass Individuum i niemals in beiden sondern nur in einem der beiden Zustände beobachtet werden kann: mit oder ohne Maßnahme. Somit können die Effekte der Maßnahme nur ermittelt werden, in dem man die realisierten Ergebnisse Y derjenigen, die an der Intervention teilgenommen haben, mit den realisierten Ergebnissen Y derjenigen, die nicht teilgenommen haben, vergleicht. In nahezu jedem Fall sind somit Datenerhebungen bei einer Kontrollgruppe notwendig. Hierbei ist es jedoch so, dass sich die Teilnehmer und Nichtteilnehmer an der Maßnahme oft in vielerlei und nicht nur in den gefragten Charakteristika unterscheiden, da eine bewusste oder unbewusste selektive Auswahl der Teilnehmer erfolgte.

Grundlage jeder Wirkungsevaluierung ist also zunächst, ein detailliertes Verständnis dafür zu entwickeln, wieso manche Individuen oder Gemeinden an der Intervention teilgenommen haben und andere nicht. Ansonsten wären Verzerrungen der Ergebnisse wahrscheinlich.

Die verschiedenen nicht-experimentellen Designs unterscheiden sich darin, wie eine Korrektur für die unterschiedlichen (beobachteten oder unbeobachteten) Charakteristika zwischen Teilnehmern und Kontrollgruppe vorgenommen werden. Hierzu sind verschiedene Annahmen erforderlich. Im Wesentlichen lassen sich hierfür vier grundlegende Herangehensweisen beschreiben, die im Folgenden genauer erläutert werden. Von diesen liefern die ersten beiden Ansätze, sofern sie anwendbar sind, oftmals überzeugendere Ergebnisse.

Instrumentalvariablenansatz (IV)

Eine Instrumentalvariable (IV) ist eine beobachtete Variable, die einen Einfluss darauf hat, ob jemand an einer Maßnahme teilnimmt oder nicht, aber zugleich nicht in Zusammenhang mit den potenziellen Ergebnissen Y_i^0 und Y_i^1 steht. Oftmals ist diese Annahme nur konditional auf weitere beobachtete Charakteristika X gültig, also wenn man nur Personen mit gleichen Charakteristika vergleicht, wie weiter unten erläutert. Es lassen sich zwei Arten von Instrumentalvariablen unterscheiden: selbst generierte und beobachtete. Die Idee ist, dass die Instrumentalvariable Z einen Einfluss auf D

und somit indirekt auf Y hat, es jedoch keinen direkten Zusammenhang zwischen beiden gibt. Eine oftmals verwendete Instrumentalvariable sind zufällige Schocks in der lokalen Niederschlagsmenge z. B. in Afrika, wo Landwirtschaft im Allgemeinen stark auf reguläre Niederschläge angewiesen ist. Zufällige Variationen im Niederschlag (*rainfall shocks*) wurden z. B. verwendet, um die langfristigen Effekte von Kinderarbeit auf Bildung und Erwerbstätigkeit im Erwachsenenalter zu schätzen. Kinderarbeit ist zweifellos eine endogene Variable, so dass eine einfache Regression von Bildung auf frühere Kinderarbeit verzerrt ist, z. B. aufgrund der Tatsache, dass reichere Familien weniger auf Kinderarbeit angewiesen sind, aber auch ohnehin mehr weiterführende Schulbildung für ihre Kinder finanziert hätten.

Zufällige kurzfristige Variationen im Niederschlag haben zehn oder fünfzehn Jahre später vermutlich keinen direkten Einfluss auf Bildung und Löhne. Andererseits zeigt sich jedoch, dass Kinderarbeit direkt auf Schocks in der Regenmenge variiert. Kurzfristige Regenschocks wirken sich auf langfristige Löhne somit nur indirekt durch Kinderarbeit aus. Dies ist die Grundidee der Schätzung der Instrumentalvariable, für die oftmals noch eine Kontrolle für weitere X Variablen, z. B. langfristige durchschnittliche Niederschlagsmenge, Urbanitätsgrad, Familienhintergrund etc. berücksichtigt werden muss.

Es besteht auch die Möglichkeit, selbst Instrumentalvariablen zu generieren, z. B. in der Form von zufällig zugeteilten Anreizen oder Ermutigungen (*random encouragement designs*). Betrachten wir z. B. den Effekt von Alkoholkonsum auf Gesundheit. Es ist vermutlich unmöglich, hier eine experimentelle Evaluation durchzuführen, welche erfordern würde, dass zufällig einige Personen keinen Alkohol konsumieren, während die zufällige Kontrollgruppe Alkohol konsumieren muss. Dies würde ja erfordern, dass die durchführende Institution Individuen zum Alkoholkonsum oder zur Abstinenz zwingen kann.

Eine simple Alternative wäre jedoch, zufällig einige Personen oder Gemeinden auszuwählen, in denen intensive Aufklärungskampagnen oder gar monetäre Anreize für die Teilnahme an Alkoholreduktionsprogrammen durchgeführt werden, während andere Personen oder Gemeinden nicht an diesen Kampagnen teilnehmen. Gleichwohl der Alkoholkonsum immer noch eine freiwillige Entscheidung der Individuen bleibt, kann die Instrumentalvariable das Selektionsproblem lösen und eine konsistente Schätzung des Effekts von Alkohol auf Gesundheit erlauben. Nach einem ähnlichen Prinzip kann z. B. der Preis für Wasser, Impfung oder Schulbücher, die Intensität der Empfehlung für Impfung oder Schulbücher etc. variiert werden. Beispielsweise können Haushalte nicht gezwungen werden, sauberes Wasser zu konsumieren. Jedoch können ihnen unterschiedliche Anreize hierzu

gegeben werden. Für die Instrumentalvariablenschätzung reicht es dann aus, zu beobachten, welches Wasser konsumiert wurde.

Diese Situation tritt nun auch in den meisten als experimentelle Studien geplanten Evaluationen auf. In experimentellen Studien treten sehr häufig zwei Probleme auf: *non-compliance* (*part-compliance*) und Befragungsausfälle (*attrition/non-response*). Gleichwohl nach ursprünglicher Planung dort Personen zufällig in Maßnahme versus Nicht-Maßnahme eingeteilt werden, gibt es oftmals Personen, die sich nicht an das experimentelle Protokoll halten. In diesem Fall kann die zufällige Zuteilung aber immer noch als Instrumentalvariable verwendet werden.

Instrumentalvariablen sind im Allgemeinen auch notwendig, um mit Problemen wie Befragungsausfällen umgehen zu können. Informationen zu den Ergebnisvariablen Y werden oftmals in Form von persönlichen Befragungen der Teilnehmer und Nichtteilnehmer durchgeführt. Hierbei wird es einerseits teilweise Antwortverweigerung geben (*non-response*), während es andererseits auch oftmals nicht möglich ist, alle Individuen wieder aufzufinden. Insbesondere zur Schätzung von mittel- und langfristigen Wirkungen möchte man Individuen nach drei, fünf oder gar zehn Jahren wiederbefragen. Aufgrund von Migration, Umzug oder Tod bzw. Krankheit ist es nicht möglich, alle Individuen wieder zu lokalisieren und zu befragen. Sofern diese Befragungsausfälle rein zufällig sind, stellen sie kein größeres Problem dar, abgesehen von Reduktionen in der Schätzgenauigkeit. Häufig sind es jedoch eher die jüngeren, mobileren, erfolgreicheren Individuen, die migrieren, so dass die verbleibende Stichprobe selektiv ist. Dieses tritt auch in experimentellen Studien auf. Instrumentalvariablen bezüglich des Interviews können helfen, eine Verzerrung der Schätzergebnisse zu reduzieren oder zu verhindern.

Regressionsdiskontinuitätsansatz (RDD)

Ein solches Design ergibt sich oftmals ungeplant in Evaluationsstudien, kann aber auch sinnvoll von vornherein in die Konzeption integriert werden. Ein üblicher Anwendungsfall findet sich, wenn die Anzahl der Maßnahmenplätze an einem Ort oder zu einem bestimmten Zeitpunkt begrenzt sind und die Auswahl der Teilnehmer nach einem objektiven Schlüssel erfolgt. Oftmals werden z. B. die Teilnehmer nach Bedürftigkeit ausgewählt, wobei Bedürftigkeit aus mehreren gemessenen Variablen berechnet wird und alle Personen oberhalb einer Bedürftigkeitsgrenze an der Maßnahme teilnehmen während alle anderen nicht daran teilnehmen. Wenn man nun alle Personen betrachtet, die entweder knapp unter der Grenze oder knapp über der Grenze lagen, so unterscheiden sich diese vermutlich wenig in ihren Charakteristika mit Ausnahme der Tatsache, dass die eine Gruppe an

der Maßnahme teilnahm und die andere nicht. Um den Grenzwert herum ist es sozusagen „zufällig“, ob eine Person knapp über oder knapp unter der Grenze landet. Wichtig ist für diesen Ansatz, dass Personen nicht individuell ihren Bedürftigkeitsindex manipulieren können, um z. B. an der Maßnahme teilzunehmen. Genaugenommen kann man eine Selektion zulassen, sofern die Individuen nicht exakt die Bedürftigkeitsgrenze kennen. Zur Plausibilitätsprüfung dieser Annahme gibt es mehrere statistische Hilfsmittel. Ein illustratives Beispiel sind Kredite oder Fördermaßnahmen für angehende Unternehmer. Diese müssen oftmals einen Geschäftsplan entwickeln, der dann von mehreren Experten bewertet wird. Nur die erfolgversprechendsten Pläne werden dann gefördert, während alle anderen nicht an der Maßnahme teilnehmen. Das Selektionsproblem wird intuitiv dadurch gelöst, dass Beobachtungen in der Nähe der Zulassungsgrenze als lokale Kontrollgruppe dienen. Obwohl das Konzept den Vorteil hat, dass es ohne eine echte Randomisierung auskommt, ist die Interpretation des geschätzten Effekts auf die Population nahe des Grenzwertes begrenzt, so dass die externe Validität der Evaluation begrenzt ist. Des Weiteren sind oftmals große Stichproben für die Schätzung erforderlich.

Propensitätsscorematching (PSM)

Im Propensitätsscorematching werden Teilnehmer und Nichtteilnehmer mit gleichen X Charakteristika verglichen, wobei die Differenzen in Y dann über die Population gemittelt werden. Wichtig ist, dass hierbei alle Variablen, die sowohl die Maßnahmenteilnahme als auch die potenziellen Ergebnisse Y_i^0 und Y_i^1 beeinflussen, in X enthalten sind. Die X Variablen sollten Variablen sein, die noch nicht von der Maßnahme beeinflusst worden sind, also vor der Maßnahme gemessen worden sein sollten. Dies erfordert daher einerseits ein sehr genaues Verständnis des Prozesses der Maßnahmenauswahl als auch eine umfangreiche Datengrundlage. Zugleich sollte man auch verstehen, wieso es vorkommt, dass Individuen mit gleichen X unterschiedliche Maßnahmen wählen. Qualitative Befragungen und Interviews mit den lokalen Projektverantwortlichen und Maßnahmenteilnehmern können hier sehr hilfreich sein, um zu erörtern, welche Variablen man in der standardisierten Befragung erheben sollte, um alle wichtigen Variablen in X aufnehmen zu können.

Die Grundidee des Verfahrens ist es, zu jedem Teilnehmer mindestens einen Nichtteilnehmer mit identischen oder sehr ähnlichen Werten von X zu finden. Die Differenz in Y zwischen diesen beiden Personen wird dann über die Population gemittelt. Eine praktische Vereinfachung besteht darin, dass es nicht erforderlich ist, dass diese Paare in ihren X Variablen identisch sind sondern dass es ausreicht, wenn sie in der Variable $p(X)$ identisch sind, die

als die Wahrscheinlichkeit der Teilnahme $p(X) = Pr(D=1|X)$ definiert ist. Es reicht somit, wenn die Balancierung der Variablen bezüglich des Propensitätsscores $p(X)$ vorgenommen wird.

Sofern Panel-Daten, d. h. Zeitreihen, bezüglich Y vorhanden sind, kann oftmals auch ein Pseudo-Treatment-Test der Plausibilität der zugrunde liegenden Annahmen durchgeführt werden. Man täuscht in den Daten vor, dass die Maßnahmenteilnehmer bereits schon vorher an der Maßnahme teilgenommen hätten und schätzt die Wirkung zu einem Zeitpunkt vor dem tatsächlichen Maßnahmenbeginn. Da die Maßnahme noch nicht stattgefunden hat, sollte die geschätzte Wirkung Null sein. Weicht die Schätzung jedoch davon ab, so ist dies ein Hinweis auf Differenzen zwischen Teilnehmern und Kontrollgruppe aufgrund von unbeobachteten Charakteristika. Sofern diese Differenzen als zeitkonstant angenommen werden können, kann ein DiD Matching Ansatz verwendet werden. Ist zu vermuten, dass die Differenzen zeitvariabel sind, so wäre dies ein Hinweis darauf, dass mehr oder bessere X Variablen oder ein besseres Verständnis des Selektionsprozesses erforderlich sind.

Differenzen in Differenzen (DiD)

Der Differenzen-in-Differenzen Ansatz kann mit oder ohne Zeitreihen und mit oder ohne Kontrolle für weitere X Variablen erfolgen. Benötigt werden hier Daten zu Y und Maßnahmenstatus D zu einem Zeitpunkt vor und nach der Einführung der Maßnahme. In der einfachsten Form wird die Differenz in Y zwischen Teilnehmern vor und nach der Maßnahme betrachtet und hiervon die Differenz genommen. Hierdurch werden zeitkonstante Unterschiede in den Charakteristika zwischen Teilnehmern und Nichtteilnehmern durch den Vergleich zu zwei verschiedenen Zeitpunkten eliminiert, so dass der reine Maßnahmeneffekt ermittelt werden kann. Das Verfahren erlaubt also eine Schätzung für unbeobachtete Differenzen zwischen Teilnehmern und Kontrollgruppe, sofern diese zeitkonstant sind.

Gleichwohl Panel-Daten oftmals hilfreich sind, ist dieses nicht strikt erforderlich. Mit anderen Worten können Querschnittsbeobachtungen vor und nach der Erhebung ausreichen. Findet die Maßnahme z. B. in der Form statt, dass bestimmte Dörfer komplett teilnehmen, andere jedoch nicht, so reicht es aus, repräsentative Befragungen in diesen Dörfern vorher und nachher durchzuführen, ohne zwingend die gleichen Individuen befragen zu müssen. Dieses kann somit auch das Problem mit Befragungsausfällen (*attrition*) bei individuellen Paneldaten umgehen.

Das einfache DiD-Verfahren eliminiert zeitkonstante Einflussfaktoren, jedoch nicht die systematischen Unterschiede in den Zeittrends zwischen Teilnehmern und Kontrollgruppe. Um letzteres Problem zu berücksichtigen,

wird man im Allgemeinen noch weitere Kontrollvariablen X in die Schätzung mit aufnehmen. Dies kann entweder in einer linearen Regression geschehen oder mittels nichtparametrischer Verfahren ähnlich zu PSM. Hier würde PSM nicht zur Identifikation sondern lediglich zur Schätzung verwendet.

Eine zusätzliche Erweiterung besteht darin, noch weitere Differenzen zu unbeeinflussten Kontrollgruppen zu nehmen. Findet eine Maßnahme z. B. nur für unter 25-jährige Anwendung, so kann die Differenz zu der doppelten Differenz zu den über 25-jährigen verwendet werden, um unterschiedliche Zeittrends, die aber zwischen den Altersgruppen gleich groß sind, zu eliminieren. Eine weitere Möglichkeit besteht, wenn Daten zu mehr als zwei Zeitpunkten vorhanden sind. Somit können differentielle Zeittrends zwischen Teilnehmer- und Kontrollgruppe eliminiert werden.

Im Gegensatz zu den nicht-experimentellen Ansätzen wird bei experimentellen Ansätzen so vorgegangen, dass man die Individuen völlig zufällig der Maßnahme oder Kontrollgruppe zuteilt, so dass sich beide Gruppen in ihren Charakteristika im Durchschnitt nicht unterscheiden. In der Praxis sind aufgrund vielfacher Komplikationen schlussendlich doch komplexere ökonometrische Verfahren notwendig.

2. Anwendungsbereiche

Quantitative nicht-experimentelle Designs lassen sich in allen Bereichen anwenden, in denen Wirkungen geschätzt werden sollen. Die hier beschriebenen Verfahren verfolgen das Ziel, die Wirkungen von definierten Maßnahmen zu messen. Sie sind allerdings nicht direkt geeignet, um die Ursachen für beobachtete Wirkungen zu suchen. Des Weiteren werden im Nachfolgenden nur mikroökonomische Verfahren beschrieben, die geeignet sind, Maßnahmen auf der Mikro- und Mesoebene zu evaluieren. Ein zentrales Element ist hier, dass es unterschiedliche Subjekte gibt, z. B. Individuen, Firmen, Krankenhäuser, Wasserwerke, Dörfer, Gemeindeverwaltungen, Stadtteile oder Bezirke etc., von denen einige von der Maßnahme betroffen sind, andere jedoch nicht. Zur quantitativen Evaluation von Projekten auf der Makroebene, z. B. Haushaltshilfen zum nationalen Budgetausgleich, sind andere ökonometrische Verfahren als die nachfolgend beschriebenen geeigneter.

Glaubwürdige nicht-experimentelle Evaluationen verlangen ein detailliertes Wissen über die Zuteilungsmechanismen sowie umfangreiche und repräsentative Daten über die Teilnehmer- und die Nichtteilnehmergruppe. Des Weiteren sind oftmals mehr als 1.000 Beobachtungen notwendig, um hinreichend präzise Schätzergebnisse über die Wirkung zu erhalten. Sie

eignen sich daher eher nicht als Monitoringinstrument für Maßnahmen in der Einführungsphase und sollten erst eingesetzt werden, nachdem die Startschwierigkeiten einer Maßnahme überwunden sind. Abgesehen von anfänglichen Hindernissen werden in den ersten Monaten oftmals noch größere Modifikationen der ursprünglich geplanten Maßnahme vorgenommen. Dies erschwert die präzise Definition der konkret untersuchten Maßnahme.

Quantitative Evaluation kann verwendet werden, um die Wirkung einer Maßnahme im Vergleich zur Situation ohne Maßnahme zu schätzen. Lehrreicher ist es jedoch, die Wirkung einer Maßnahme im Vergleich zu anderen Maßnahmen zu schätzen bzw. zu verschiedenen Varianten der Maßnahme zu schätzen, bei gleichem kontextuellen Hintergrund und mit der gleichen Datenerhebung. Somit kann z. B. betrachtet werden, welche Anreize oder Kostenbeteiligungsstrukturen für subventionierte Ausbildung im handwerklichen Gewerbe für welche Individuen erfolgreicher sind.

3. Analyseansatz

Ziel der Evaluation ist eine möglichst glaubwürdige Schätzung der Wirkungen der Maßnahme. Die Glaubwürdigkeit der Studie hängt von der Plausibilität der zugrunde liegenden Annahmen ab. Gleichwohl die ökonometrischen Schätzungen oftmals erst nach Abschluss der Maßnahme sowie der Datenerhebung stattfinden, basiert die Evaluation auf Annahmen über die Einführung der Maßnahme. Daher muss unterschieden werden zwischen

- a) Evaluationsansätzen ex-ante (also vor Einführung der Maßnahme);
- b) Identifikationsstrategien ex-post (nach Durchführung der Datenerhebung);
- c) ökonometrischen Schätzverfahren.

Prinzipiell werden Wirkungsschätzungen in mehreren Dimensionen erstellt. Dies erfolgt einerseits für verschiedene Ergebnisvariablen, oftmals geordnet nach vermuteten Wirkungsketten. Beispiele sind der Effekt eines Mikrokreditprogramms für Kleinunternehmer auf die Bilanzsumme, Investitionen, Umsatz, Profit, Einstellung weiterer Mitarbeiter, Gesundheitszustand des Unternehmers, Schulbesuch seiner Kinder, Rückzahlung des Kredits, Eröffnung weiterer Kreditzugangskanäle etc., unterschieden nach kurz-, mittel- und langfristigen Effekten. Zweitens werden geschätzte Effekte üblicherweise für verschiedene Untergruppen angegeben, z. B. nach Geschlecht, Alter, Wohnregion etc.

Zum Dritten ist es interessant, die Wirkungen verschiedener alternativer Maßnahmen untereinander zu vergleichen, z. B. eines Mikrokreditprogramms mit Gruppenhaftung gegenüber individueller Haftung.

Bei der Interpretation der Schätzergebnisse sollte berücksichtigt werden, dass nicht-experimentelle Ansätze im Allgemeinen wesentlich unpräzisere Schätzungen liefern als experimentelle Ansätze. Falls die Konfidenzintervalle sehr groß ausfallen, sollten statistisch nicht-signifikante Ergebnisse nicht als Evidenz für die Abwesenheit einer Wirkung interpretiert werden! Diese Interpretation ist nur zulässig, wenn die Konfidenzintervalle sehr eng sind. Die korrekte Interpretation wäre, dass die Stichprobengröße zu klein war, um zuverlässige Aussagen zu treffen.

4. Durchführung (inkl. Datenerhebung und Datenauswertung)

a) Verfahrensablauf

Projektplanung

Die Glaubwürdigkeit der Evaluation kann oftmals deutlich erhöht werden, wenn bereits vor der Einführung der Maßnahme über die später vorzunehmende Evaluation nachgedacht wird. Dies betrifft im Wesentlichen zwei Elemente:

Das wichtigste, und oftmals am wenigsten beachtete, ist die Einführung von Mechanismen zur Erzeugung exogener Variation in der Teilnahmenscheidung, wie es oben bereits dargelegt wurde. Denn bei der späteren Evaluation wird sich stets die Frage stellen, wieso manche Individuen (mit identischen Charakteristika) an einer Maßnahme teilgenommen haben und andere nicht oder wieso manche Dörfer teilgenommen haben und andere nicht. Sofern es hierzu keine überzeugende Erklärung gibt, wird sich stets die Vermutung aufdrängen, dass dies auf unterschiedliche unbeobachtete Charakteristika zurück zu führen sei, welche dann die Glaubwürdigkeit der Evaluationsergebnisse in Frage stellt. Oftmals konnten aus Budgetgründen nur wenige Dörfer teilnehmen. Diese wurden jedoch vermutlich nicht vollkommen zufällig sondern nach Kriterien wie Anbindung an das Straßennetz oder nach Bedürftigkeit ausgewählt, so dass sich Teilnehmer und Kontrollgruppe in wichtigen Charakteristika unterscheiden. Ähnlich sieht es mit der Teilnahme von einzelnen Personen an Maßnahmen aus, welche tendenziell eher besser über das Programm informiert sind oder sich einen höheren Nutzen von der Teilnahme versprechen. Diese unterscheiden sich somit, zumindest im Durchschnitt, in wichtigen Eigenschaften von den Nichtteilnehmern.

Die wichtigsten Wege, dieses Problem zu vermeiden, sind – sofern die völlig zufällige Zuteilung von Individuen zu Teilnehmer- und Kontrollgruppe (= experimenteller Ansatz) nicht möglich ist, die oben dargelegten

Ansätze mit generierten Instrumentalvariablen oder durch Einbau einer Regressionsdiskontinuität. Sofern keiner dieser beiden Ansätze möglich ist, sollten einerseits detaillierte Informationen über Teilnehmer und Nichtteilnehmer gesammelt werden und andererseits während des Projekts sehr genau dokumentiert werden, nach welchen Kriterien und Prozeduren Teilnehmer ausgewählt wurden.

Wichtig sind zudem eine sinnvolle Definition des Beobachtungsobjekts und eine präzise Definition des Zustands der Nichtteilnahme. Die Beobachtungseinheit sollte möglichst so gewählt werden, dass Nichtteilnehmer wirklich nicht von der Maßnahme beeinflusst worden sein können. Ist z. B. eine völlig zufällige Zuteilung von Schulbüchern möglich, so sollte dieses möglichst auf der Ebene der Schulen geschehen und nicht auf der Ebene der Schüler, da Schüler, die zwar selbst kein Schulbuch erhalten haben, dennoch von den Schulbüchern ihrer Mitschüler profitieren können (= Externalitäten). Schüler ohne Schulbuch wären also trotzdem von der Maßnahme betroffen falls einige ihrer Mitschüler Bücher erhielten.

Des Weiteren sollte bereits in der Projektvorbereitung die Erhebung von Daten über Teilnehmer und Kontrollgruppe geplant werden. Daten sind erforderlich sowohl über die interessierenden Ergebnisvariablen nach der Maßnahme – idealerweise zu verschiedenen Zeitpunkten, um kurzfristige und längerfristige Wirkungen schätzen zu können – sowie Informationen über wichtige Charakteristika vor der Teilnahme oder Anmeldung zur Maßnahme. Zusätzlich sind Informationen zum Teilnahmestatus und während der Maßnahme erforderlich. Letzteres ist insbesondere deshalb wichtig, da vermeintliche Nichtteilnehmer eventuell an ähnlichen oder identischen Maßnahmen (z. B. einer anderen Organisation) teilgenommen haben könnten. Individuen denen z. B. die Zuteilung eines Ausbildungskredits verweigert wurde haben eventuell einen ähnlichen Kredit im Nachbarort erhalten.

Die Vorbereitung der Datenerhebung erfordert die Definition der zugrunde liegenden Grundgesamtheit, Verfahren zur Auswahl repräsentativer Stichproben (sofern keine Vollerhebung möglich ist) sowie der Befragungstechniken. Wichtig ist, dass die gleichen Befragungsinstrumente für Teilnehmer und Kontrollgruppe verwendet werden.

Es sollte auch geprüft werden, ob eventuell auf bereits bestehenden Datensätzen, z. B. Haushaltsbefragungen, aufgebaut werden kann.

Des Weiteren finden einige der Aspekte zur experimentellen Evaluationshier ebenfalls Anwendung. Für die Entwicklung der Datenerhebungsinstrumente sollten die Variablen identifiziert werden, auf die das Projekt mutmaßlich einen Einfluss haben könnte. Hierzu ist die Erstellung hypothetischer Wirkungsketten hilfreich, um sicherzustellen, dass alle wichtigen Variablen auch in der Befragung erhoben werden. Diese Überlegungen sind

wichtig, um spätere Zuordnungslücken zu vermeiden. Die Datenerhebungsinstrumente können jedoch auch für die zweite oder dritte Nachbefragung noch ergänzt werden.

Baseline-Befragung

Eine Befragung von Teilnehmern und Kontrollgruppe vor dem Beginn der Maßnahme ist nicht zwingend für alle Verfahren erforderlich aber hilfreich, zumindest um Konsistenzprüfungen durchführen zu können oder um spätere Befragungsausfälle (*attrition, non-response*) bei Folgebefragungen korrigieren zu können. Für einige Verfahren wie z. B. Propensitätsscorematching (PSM) ist eine *baseline*-Befragung nahezu immer erforderlich, um glaubwürdige Evaluationsergebnisse zu erhalten. Idealerweise sollten bei der *baseline*-Befragung möglichst auch alle später zu verwendenden Ergebnisvariablen erhoben werden, um einen Pseudotreatmenttest zu ermöglichen.

Durchführung der Maßnahme

Während der Maßnahme erfolgen begleitende Datenerhebungen z. B. über die Zeitpunkte wann ein Individuum sich zur Maßnahmenteilnahme gemeldet hat, wann die Maßnahme begonnen hat, wann die Maßnahme endete etc. Wichtig ist es auch sicherzustellen, dass das Protokoll zur Erzeugung exogener Variation eingehalten und relevante Informationen hierzu gesammelt werden. Wird zum Beispiel ein Regressionsdiskontinuitätsansatz verwendet, bei dem zuerst die Bedürftigsten in eine Maßnahme aufgenommen werden, so ist zu prüfen, ob die Bestimmung der Bedürftigkeit korrekt anhand von beobachteten Charakteristika ermittelt wurde.

Erste Befragung nach Abschluss der Maßnahme

Zur Bestimmung der kurzfristigen Wirkung der Maßnahme wird nach ca. einem halben bis einem Jahr eine Befragung der Teilnehmer und der Kontrollgruppe vorgenommen. Sofern nicht mit einem vollen Rücklauf gerechnet werden kann, sollten Elemente für Instrumentalvariablen bezüglich des Rücklaufs eingebaut werden, z. B. Charakteristika über die Interviewer gesammelt werden.

Weitere Befragungen

Zur Einschätzung der Nachhaltigkeit einer Maßnahme sollten weitere Befragungen der Teilnehmer und der Kontrollgruppe nach mehreren Jahren vorgesehen werden. Dies verlangt intensive Vorbereitungen, um sicherzu-

stellen, dass möglichst viele Teilnehmer und Nichtteilnehmer auch nach mehreren Jahren noch aufgefunden und zur Teilnahme an der Befragung ermuntert werden können.

Ökonometrische Auswertung

Die ökonometrische Auswertung besteht aus drei Teilen. Im ersten Schritt werden die Daten aufbereitet, verifiziert und auf Konsistenz geprüft. Im nächsten Schritt wird die Identifikationsstrategie festgelegt, unter Berücksichtigung des geplanten Evaluationsansatzes sowie der bei der Projektdurchführung und Datenerhebung aufgetretenen Schwierigkeiten. War z. B. eine experimentelle Evaluation geplant, so zeigen sich in der Praxis oftmals doch sehr viele Schwierigkeiten, die schlussendlich doch komplexere ökonometrische Verfahren erfordern. Typische Probleme sind Abweichungen vom experimentellen Protokoll in der Form, dass Individuen, die zur Maßnahme eingeteilt wurden, nicht an dieser teilgenommen haben und/oder Individuen, die zur Nichtteilnahme eingeteilt wurden, trotzdem Zugang zur Maßnahme erhielten. Ein weiteres Problem ist oft, dass Individuen nur partiell an einer Maßnahme teilgenommen haben. Des weiteren treten oft Probleme bei der Datenerhebung auf, z. B. in der Form, dass die Stichprobe nicht wirklich repräsentativ erhoben wurde oder dass Individuen nicht aufgefunden werden konnten, Antworten in der Erhebung ganz oder teilweise verweigert wurden oder inkonsistent waren. Letztere Probleme nehmen oftmals zu, wenn die Befragung zur Messung von mittel- und langfristigen Effekten wiederholt wird. Im letzten Schritt erfolgt die ökonometrische Schätzung.

b, c, d) Akteure, Datenerhebung und -auswertung

Prinzipiell könnte die gesamte Evaluation durch das Projektteam durchgeführt werden. Es bietet sich jedoch aus zwei Gründen an, das Projekt und dessen Evaluation zu trennen. Einerseits wird dadurch die Objektivität der Evaluation gestärkt. Zum Zweiten verfügt das Projektteam, das sich auf die Organisation der Maßnahme konzentriert, oftmals nicht über spezifisches Wissen in Stichprobenauswahl, Datenerhebung und ökonometrischen Analyseverfahren. Dennoch sollte eine Zusammenarbeit zur Vorbereitung der Evaluation erfolgen, möglichst bereits schon während der Projektplanung, wie oben beschrieben wurde.

Für die Datenerhebung vor Ort ist die Hinzunahme eines lokalen Befragungsinstituts hilfreich bzw. die Rekrutierung bereits ausgebildeter und zuverlässiger Interviewer.

5. Nutzung der Ergebnisse

Ähnlich wie bei experimentellen Methoden sollten die Evaluationsergebnisse einer möglichst breiten Öffentlichkeit verfügbar gemacht werden und idealerweise mit Evaluationsstudien in anderen Ländern oder mit ähnlichen Maßnahmen vergleichbar gemacht werden. Somit kann nicht nur die durchführende Organisation sondern können auch andere Organisationen etwas über die Wirksamkeit der Maßnahmen erfahren, um sukzessive mittels wiederholter Evaluationsstudien Wissen zu akkumulieren. Hierzu ist eine möglichst standardisierte Darstellung der Ergebnisse wichtig. Dies verlangt neben einer genauen Beschreibung der Maßnahme auch Angaben über geschätzte Konfidenzintervalle, Stichprobengrößen und die verwendeten ökonometrischen Verfahren sowie eine Einschätzung der Zuverlässigkeit der geschätzten Ergebnisse. Selbst im Idealfall einer experimentellen Evaluation treten in der Praxis im Laufe des Projekts oder der Datenerhebung oftmals Probleme auf, welche die grundlegenden Annahmen der Randomisierung verletzen können. Während das Konfidenzintervall die statistische Schätzunsicherheit widerspiegelt, unter der Prämisse, dass alle identifizierenden Annahmen korrekt sind, ist eine Einschätzung der Zuverlässigkeit der identifizierenden Annahmen eher subjektiv. Für den Leser mehrerer Evaluationsstudien ist es jedoch wichtig abschätzen zu können, welchen Evaluationsergebnissen eher weniger und welchen eher mehr zu trauen ist.

6. Anforderungen

Die Anwendung nicht-experimenteller Designs ist tendenziell weniger aufwändig als die experimenteller Ansätze. Insbesondere ist mit weniger Kritik auch bezüglich ethischer Bedenken zu rechnen. Auf der anderen Seite ist die ökonometrische Analyse selbst dafür eher komplexer und verlangt Fachlichkeit, wie die oben stehenden Erläuterungen zu den Ansätzen deutlich machen.

Real Time Evaluation (RTE)

CLAUDIA CONRAD

1. Allgemeine Informationen

Real Time Evaluations (RTE, Echtzeit-Evaluierungen) finden ihren Ursprung in den verschiedenen Bereichen der technologischen (vorwiegend amerikanischen) Forschung sowie der neueren IT-Anwendung. Sie werden (heute) v. a. im Rahmen der humanitären Hilfe durchgeführt. Eine erste Referenz zur Anwendung von RTE in der humanitären Hilfe findet man in Alistair Hallam's ‚Good Practice Review Evaluating Humanitarian Assistance Programmes in Complex Emergencies‘. Die neunziger Jahre waren das Jahrzehnt, in dem umfangreiche Interventionen im Rahmen der Humanitären Hilfe geplant und durchgeführt wurden (Ruanda 1994, Balkan-Krieg ab 1992). Zahlreiche staatliche und nicht-staatliche Organisationen machten es sich zur Aufgabe, der notleidenden Bevölkerung zu helfen – darunter viele Organisationen, die vormals nur reine Entwicklungsprojekte durchführten. RTE wurde als ein internes Korrektiv eingeführt, um die Effektivität und den direkten Nutzen der laufenden Interventionen gewährleisten zu können.

Im humanitären Bereich ist v. a. UNHCR mit seiner Evaluierungsstelle (Evaluation and Policy Analysis Unit – EPAU) der wichtigste und größte Anwender von RTE und publiziert die Berichte. Auf der Homepage findet man unter den FAQ alles Wissenswerte (www.unhcr.org/research/research/3ce372204.pdf). WFP, UNICEF, the Humanitarian Accountability Project, CARE, World Vision, OXFAM GB, die Föderation des Internationalen Roten Kreuzes, FAO und andere haben ebenfalls die regelmäßige Durchführung von RTE eingeführt. Das britische Disaster Emergency Committee (DEC) führt vergleichbare Prozesse unter dem Namen ‚Monitoring Mission‘ durch. Die Groupe Urgence Réhabilitation Développement (URD) gestaltet RTE in einem iterativen Evaluierungsprozess (l'évaluation itérative avec mini-séminaires – EIMS, vgl. www.urd.org).

Die Nachfrage an methodischer Begleitung während der Durchführung einer Operation in insbesondere komplexen Nothilfesituationen stieg mit den Erfahrungen aus Ruanda wesentlich an. Es sollte gewährleistet werden, dass möglichst zahlreiche Menschen effizient (v. a. schnell) und effektiv

durch die Interventionen erreicht werden. Evaluierungen, die üblicherweise im Anschluss eines Nothilfe-Programms stattfinden (Ex-post-Evaluierungen), erwiesen sich als ungenügend, um rechtzeitig Einfluss auf die Durchführung von humanitären Interventionen nehmen zu können (Leben konnten nicht mehr gerettet werden, Korrekturen fanden nicht mehr dort statt, wo sie nötig gewesen wären, die Interventionen waren meist schon lange eingestellt, die involvierten Mitarbeiter längst an anderer Stelle eingesetzt, das Geld bereits verbraucht). Die Tatsache, dass bei den meisten Organisationen das institutionelle Gedächtnis eher schwach ausgeprägt ist, nährte zudem eine gewisse Skepsis, ob die gemachten Erfahrungen wirklich dazu beitragen können, zukünftige Interventionen angemessener zu gestalten.

2. Anwendungsbereiche

RTE werden in erster Linie im Rahmen der humanitären Hilfe, insbesondere in komplexen Notsituationen (z. B. Darfur/Sudan) angewandt. Ein Schlüsselmerkmal von RTE ist die Tatsache, dass sie einen direkten Einfluss auf die laufende Programmierung von Interventionen hat. Dies rückt RTE konzeptionell in die Nähe von Monitoring-Aufgaben und fordert damit die klassische Unterscheidung zwischen ‚Monitoring‘ und ‚Evaluierung‘ heraus.

In einem laufenden Programm können mehrere RTE durchgeführt werden, RTE sind integraler Bestandteil des *Project Cycle Management* (PCM). Dazu benötigt es die aktive Unterstützung und Initiierung der Durchführung durch das Management. Diskussionen über Programmqualität werden üblicherweise geführt, erscheinen jedoch nicht unbedingt in den Berichten.

3. Analyseansatz

Allgemein beinhalten RTE die Überprüfung der Effektivität und Effizienz einer Intervention sowie des intendierten Nutzens für die Zielgruppe. Oftmals bewerten RTE sowohl Interventionen an sich als auch die Handlungsstrategien, die diesen Interventionen zugrunde liegen. Die Bewertungen konzentrieren sich darauf, wie die Ergebnisse erreicht werden (der Schwerpunkt liegt auf der Prozessebene statt auf der Wirkungsebene). In Anbetracht der Tatsache, dass RTE einen sehr begrenzten Zeithorizont umfassen, sind begründete Aussagen zu intendierten oder nicht-intendierten Wirkungen auch nicht unbedingt zu erwarten. Trotzdem können die Ergebnisse einer RTE durchaus im Sinne einer beschreibenden Grundlagenherhebung

(*baseline*) in spätere Evaluierungen, welche auf die Wirkungserfassung ausgerichtet sind, einfließen.

RTE soll es ermöglichen, humanitäre Interventionen auf der Basis von festgelegten *benchmarks* zu bewerten. Dazu gehören die jeweiligen institutionellen Politikgrundsätze (auch Leitbilder), Mandatsbeschreibungen, sektorale oder thematische Leitlinien, Humanitäre Standards sowie die spezifischen Zielsetzungen der Intervention.

RTE sind in erster Linie so angelegt, dass sie auf eine unmittelbare Verbesserung der Programmimplementierung abzielen (*improvement oriented review*). RTE machen nur beschränkten Gebrauch der DAC-Kriterien für Evaluierung; ihr Fokus liegt eher auf dem Prozess der Durchführung. Die Relevanz einer Intervention wird allenthalben mit den für die humanitäre Hilfe geltenden Standards wie den Code of Conduct der Rotkreuz-Gesellschaften sowie den Sphere-Standards gemessen.

4. Durchführung

a) Verfahrensablauf

Eine RTE findet im laufenden Programm so früh wie möglich statt. In der Praxis heißt dies: die Durchführung einer RTE wird bereits in der Planung der humanitären Intervention terminiert. Idealerweise sind die Evaluatoren nicht nur bereits bestimmt, sondern wirken bei der Planung der Interventionen mit. RTE sind zudem schnell und dauern nur wenige Tage. Die Berichterstattung erfolgt innerhalb 4-5 Tagen nach Ende der RTE-Mission. Eine RTE gliedert sich in drei Phasen:

Vorbereitung

Es werden Themen zur Programmierung und zum Management identifiziert und aufgelistet. Eine Leistungsbeschreibung i. S. v. Terms of Reference gibt es nicht. Je nach Thema arbeitet der Evaluator hauptsächlich in der Zentrale oder im Interventionsgebiet. In den UN-Organisationen nimmt der RTE-Evaluator einen beratenden Platz im Krisenstab für die jeweilige Nothilfe-Interventionen ein. Bereits in der Vorbereitung können erste Interviews mit dem Führungspersonal durchgeführt werden. Es wird erwartet, dass der RTE-Evaluator bereits in dieser frühen Phase beratend den zuständigen Mitarbeitern zur Seite steht.

Durchführung

Die Durchführung einer RTE im Feld muss nicht zwingend mit der Lancierung der ersten Nothilfe-Maßnahmen einhergehen. Üblicherweise wird das Evaluatorenteam vier bis sechs Wochen nach Interventionsbeginn in das Feld geschickt. Je nach thematischem Aufgabenbereich wird der Einsatz des einzelnen Evaluators gestaltet. Bei Abschluss der RTE wird im Feld mit den betroffenen Mitarbeitern ein umfassendes *de-briefing* durchgeführt, an dem durchaus auch Vertreter anderer Partnerorganisationen teilnehmen können. Während der Fokus von ‚konventionellen‘ Evaluierungen eher auf der Datenerhebung im Feld und anschließender Berichterstattung mit Empfehlungen zuhanden des Auftraggebers liegt, sind RTE in erster Linie ‚Unterstützungs-Dienstleister‘ in der laufenden Intervention.

Abschluss

Eine RTE wird üblicherweise mit einem *de-briefing* in der Zentrale sowie der zügigen Vorlage des Berichtes abgeschlossen. Der Bericht ist eher eine Momentaufnahme zum Zeitpunkt der Evaluierung. Daher können sich nach Vorlage des Berichtes die Bedingungen bereits erheblich geändert haben. Der Bericht sollte Vorschläge zur Fortführung weiterer RTE anhand konkret spezifizierter Themen enthalten.

b) Akteure

Das Evaluatorenteam einer RTE setzt sich aus internen und externen Personen zusammen. Letztere müssen zwingend der Organisation nahe stehen und diese gut kennen, da die zur Verfügung stehende Zeit nicht ausreichend ist, um die internen Prozesse der Organisation hinreichend zu erfassen und zu bewerten.

Der Evaluator in einer RTE ist eher mit einem *facilitator* vergleichbar, der den kritischen Blick auf das Geschehen wirft und in erster Linie kreative Lösungen erarbeiten soll. Ebenfalls kann der Evaluator unterstützend begleiten, wenn eine Organisation ein besseres Verständnis eines bestimmten Sachverhalts erreichen möchte (so geschehen in Darfur-Interventionen).

Grundsätzlich wirkt der Evaluator unterstützend dem Programmpersonal gegenüber (Erleichterung von Stresssituationen). Dies gilt insbesondere jungen oder neuen Mitarbeitern gegenüber. In der humanitären Hilfe sind Personalwechsel relativ häufig, da nicht alle Mitarbeiter beliebig lang unter Hochdruck arbeiten können. Diese Form der Unterstützung darf nicht als ‚weiches Nebenprodukt‘ einer RTE angesehen werden, sondern stellt ein wichtiges und zentrales Element in dieser Evaluierungsform dar.

Idealerweise kennt der Evaluator allgemeine Schlussfolgerungen (*lessons learned*) aus vergleichbaren anderen humanitären Situationen und ist in der Lage, dieses Wissen schnell und präzise auf die spezifische Situation zu übertragen und anzuwenden. Ziel ist es, Lösungen fristgerecht zu erarbeiten, um die effektive Durchführung der Intervention zu unterstützen.

Die RTE findet in einem iterativen Prozess statt. Zentrales Element ist der starke Einbezug der Akteure und der Austausch zwischen den Akteuren in der Zentrale (*headquarter*) und dem Feld (Interventionsgebiet).

c) Datenerhebung

Für die Datenerhebung werden konventionelle Methoden wie Stakeholder-Interviews, Befragung der Zielgruppen, Beobachtungen, Analyse von Programmdokumenten und Sekundärliteratur, etc. angewandt. Dabei arbeitet RTE meist ohne Stichprobenverfahren, da Schnelligkeit ein wichtiges Kriterium der Durchführung ist und der Fokus auf der beratenden Leistung des operativen Teams liegt.

RTEs können sowohl in der Zentrale als auch im Interventionsgebiet durchgeführt werden. Es ist sogar aufgrund der Fristigkeit und eventueller Zugangsbeschränkungen zu entsprechenden Akteursgruppen üblich, Interviews per Telefon zu führen.

d) Datenauswertung

Das Evaluatorenteam interpretiert die Daten entlang organisationsspezifischer Standards und Standards der humanitären Hilfe. Das Team spiegelt die Ergebnisse unmittelbar der Organisation zurück (respektive dem Einsatzteam vor Ort) und ermöglicht so direkte Änderungen/Anpassungen auf Implementierungsebene.

5. Nutzung der Ergebnisse

Die Ergebnisse der RTE fließen unmittelbar in die laufende humanitäre Intervention ein und können deren Verlauf erheblich beeinflussen oder sogar grundsätzlich ändern. Damit kommt RTE eine Steuerungsfunktion zu. Organisationen, die RTE regelmäßig durchführen, schätzen folgenden Nutzen: Die Evaluierung ist frist- und zeitgerecht entlang der laufenden Intervention und bietet so die Möglichkeit zur kritischen Reflektion. Das Feedback ist sehr schnell vorhanden – ein Muss in akuten Nothilfesituationen. Die Mitarbeiter schätzen beim Evaluationsteam die Kombination zwischen der in-

stitutionellen Nähe (kein zeitaufwändiger Erklärungsbedarf) und einer kritischen Distanz, da die Evaluatoren nicht im operativen Geschäft verhaftet sind.

Da RTE eine Grundlage für (zumeist strategische) Entscheidungen darstellt, wird die Nachvollziehbarkeit der Ergebnisse und Bewertungen und ihre Transparenz für die Beteiligten zwangsläufig berücksichtigt. Quantifizierungen der Daten können bei RTE eine sehr große Rolle spielen, wenn es z. B. um Verteilungsfragen geht und die Zielerreichung in Frage gestellt ist oder aber die *coverage* – also die Reichweite der Intervention – bewertet werden soll.

Der inhaltliche Schwerpunkt von RTE wird auf die Erfassung von *lessons learned* gelegt. Aspekte der Rechenschaftslegung treten demgegenüber zurück.

6. Anforderungen

Der zeitliche Aufwand für die Vorbereitung einer RTE beträgt 2-5 Tage, für die Durchführung 4-5 Tage, für die Berichterstattung maximal 5 Tage.

Der Evaluator muss die Organisation und deren Prozesse gut kennen. Zudem verlangt die Durchführung einer RTE ein hohes Maß an Flexibilität (Begleitung von Soforteinsätzen) und sehr gute Kenntnisse der humanitären Hilfe.

7. Besondere Eignung und Grenzen des Verfahrens

RTE sind durch die enge Anbindung an das Programm-Management (im Feld und/oder in der Zentrale) ein guter Katalysator für a) institutionelles Lernen und b) die Gewährleistung von qualitätssichernden Maßnahmen. Aus der Sicht von Befürwortern rigoroser Evaluierungsmethoden weist RTE gewisse Inkonsistenzen in der Methodologie auf, ist theoretisch zu wenig untermauert und besitzt einen zu geringen Standardisierungsgrad.

Literaturhinweise

Hallam, Alistair (1998): Good Practice Review Evaluating Humanitarian Assistance Programmes in Complex Emergencies' (Hrsg. Humanitarian Practice Network, 1998).

Wissenswertes auf der UNHCR-Homepage:

www.unhcr.org/research/RESEARCH/3ce372204.pdf

Weitere Artikel zum Einstieg und zur praktischen Anwendung:

Maurice Herson “Real – time evaluation – where does its value lie?” (<http://www.odihpn.org/report.asp?ID=2772>).

Maurice Herson „Real-time evaluations in Darfur: Some suggestions for learning” (<http://www.odihpn.org/report.asp?id=2526>).

P. Sandison “2003 Global: Desk Review of Real-time Evaluation experience” (http://www.unicef.org/evaldatabase/index_29680.html).

IFRC “Real-time evaluation of tsunami response in Asia and East Africa” (www.ifrc.org/docs/pubs/updates/rte2sythesis.pdf).

Verbleibsstudien (Tracer Studies)

BERNWARD CAUSEMANN / RALF LANGE

1. Allgemeine Informationen

Absolventenverbleibsstudien (englisch: tracer studies) sind ein lange erprobtes Mittel in der Evaluierung und Selbstevaluierung von formalen und non-formalen Berufsbildungseinrichtungen, Stipendienprogrammen und Maßnahmen der Beschäftigungsförderung. In diesem Steckbrief werden sie mit besonderem Fokus auf der Anwendung für Selbstevaluierungen dargestellt, aber auch Bezüge zur Anwendung bei externen Evaluierungen hergestellt.

FAKT hat bei vielen Partnerorganisationen von Brot für die Welt und EED bzw. EZE Verbleibsstudien durchgeführt und sie beraten, wie sie das Instrument für Selbstevaluierung anwenden. In Vorhaben der GTZ wurden Verbleibsstudien als Monitoringmethode eingesetzt. Ziel ist bei Selbstevaluierung, dass die Einrichtungen tertiärer und beruflicher Bildung und Träger von Beschäftigungsfördermaßnahmen aus den Ergebnissen für die weitere Entwicklung ihrer Aus-, Fortbildungs- und Arbeitsvermittlungsangebote lernen. Verbleibsstudien sollen ein zentraler Teil des Management-Informationssystems sein. Verbleibsstudien finden auch Anwendung im Vorfeld der Evaluierung von Berufsbildungs- und Stipendienprogrammen. Ziel externer Evaluierungen ist es, Aufschlüsse über die Wirkung von Ausbildungen zu bekommen.

2. Anwendungsbereiche

Verbleibsstudien werden in den Sektoren berufliche Bildung, Beschäftigungsförderung und Hochschulbildung angewandt. Sie werden eingesetzt sowohl bei längerfristigen Trainings und Ausbildungen wie bei Kurzzeitmaßnahmen.

Verbleibsstudien sind auf der Mikroebene einsetzbar z. B. bei Berufsbildungseinrichtungen; auf der Mesoebene in Programmen, welche Wirkungen auf regionaler Ebene anstreben (z. B. regionale Ansätze der Arbeitsvermittlung, regionale Beschäftigungsförderprogramme, lokale öko-

nomische Entwicklung); auf der Makroebene im Rahmen von Sektorstudien und beim Monitoring von Vorhaben der Bildungsreform.

Bezogen auf die einzelnen Teilnehmerpopulationen ist der Einsatz *ex-post*, bezogen auf ein längerfristiges Bildungsprogramm können Verbleibsstudien im Verlauf und *ex-post* eingesetzt werden. Sie erfassen Lernerfolge und deren Anwendung.

3. Analyseansatz

Das Instrument lässt viel Spielraum, welche Wirkungen erfasst werden sollen. Geplante wie ungeplante Wirkungen können abgefragt werden. Über die direkte Nutzung von Outputs hinaus werden viele weitere indirekte und schwer zuordenbare Wirkungen beschrieben, aber auch ausbleibende Wirkungen, wenn die Ausbildung vom Markt nicht nachgefragt wird, und Marktsättigung, wenn zu viele Personen in einer Fachrichtung ausgebildet werden.

Verbleibsstudien befragen diejenigen, die an Ausbildungs- und Bildungsmaßnahmen teilgenommen haben, nach ihrer aktuellen Situation und nach ihrer Einschätzung, wie die Maßnahme bei ihnen gewirkt habe. Dies geschieht in Kenntnis der Kompetenz der Teilnehmer vor Beginn der Maßnahme, so dass ein Vorher-nachher-Vergleich gemacht werden kann.

Ob Verbleibsstudien mit dem Zuordnungsproblem implizit oder explizit umgehen, hängt vom eingesetzten Fragebogen ab, der jeweils an die Situation angepasst wird. Die Zuordnung der Wirkung geht von subjektiver Einschätzung aus und muss dann in der Analysephase und durch andere Instrumente überprüft werden.

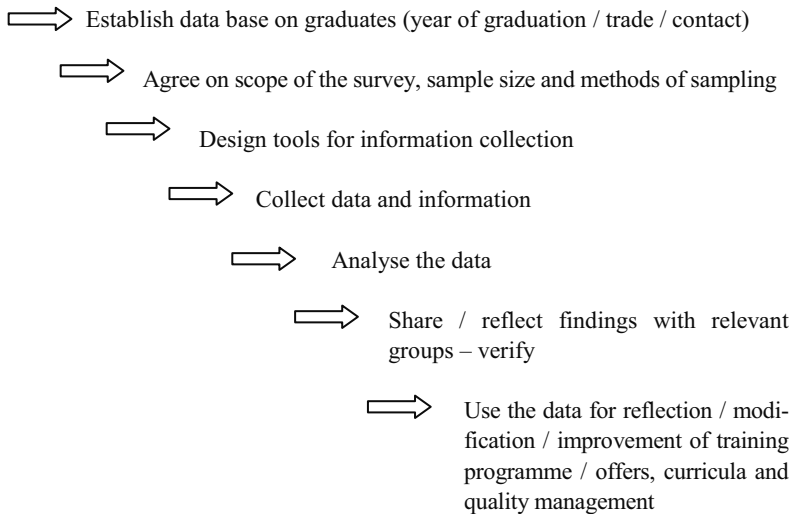
4. Durchführung

a) Verfahrensablauf

Für Verbleibsstudien wird zunächst erhoben, welche Absolventen es gab, die Größe des *samples* bestimmt (Gesamtbefragung einer Teilpopulation, Zufallsauswahl oder Selektion nach festzulegenden Kriterien) und ein Befragungsinstrument entwickelt, das i. d. R. schriftlich angewandt wird. Mit diesem Instrument werden Daten erhoben und diese Daten analysiert. Dazu sollte der Vergleich mit den Ausgangsdaten gehören. Ergebnisse der Analyse werden dann mit relevanten Beteiligten oder Interessierten reflektiert und

die Erkenntnisse wie die Schlussfolgerungen verifiziert. Schließlich werden die Daten für die Reflexion und Modifikation des Programms genutzt.

ABBILDUNG 16: Verbleibsstudien integriert in MIS von Berufsbildungseinrichtungen



Nach: Lange, 2001, S. 2

b) Akteure

Voraussetzung für die Durchführung von Verbleibsstudien im Vorfeld von Evaluierungen und als Instrument der Selbstevaluierung ist die Sensibilisierung von Entscheidungsträgern (Führungskräften) in den betreffenden Berufsbildungseinrichtungen. Bei mangelnder Kapazität ist externe Begleitung notwendig. Ausbilder sollten in die Auswertung der Ergebnisse einbezogen werden, aber nicht unbedingt in die Datenerhebung (Gefahr der Verzerrung von Ergebnissen). Im Idealfall verfügt der Träger über Strukturen der Nachverfolgung. Das darin eingesetzte Personal ist besonders für die Durchführung von Verbleibsstudien geeignet. Insbesondere bei nicht-akademischen Zielgruppen ist eine hohe Rücklaufquote nur zu erreichen, wenn die Bildungseinrichtung/Beschäftigungsförderinitiative über ein System des Nachkontakts verfügt. Hohe Mobilität von Absolventen (Arbeitsmigration) erschwert die Durchführung. Die Absolventen selbst sind eine wichtige Akteursgruppe und sollten über Fokusgruppen in den Verifizierungsprozess

eingebunden werden. Verbleibsinformationen können indirekt über Befragungen von Arbeitgebern erhoben werden.

Bei Verwendung des Instruments im Management-Informationssystem einer Bildungseinrichtung dient es dazu, die Kenntnis aller Beteiligten über das externe Umfeld und die Wirksamkeit der Dienstleistungsangebote/Projektinterventionen zu verbessern und die Institution entsprechend weiterzuentwickeln. Bei einmaliger Verwendung in externer Evaluation ist die Prozessorientierung nicht zwingend.

c) Datenerhebung

Datenquellen sind vor allem Wissen und Einschätzungen der Absolventen und indirekt von Arbeitgebern. Das *sampling* ist von den Evaluatoren oder der Institution frei wählbar. Eine Zufallsauswahl wird angeraten. Für einen Trainingszyklus hat das Instrument den Charakter einer Evaluierung, für die Institution den Charakter fortlaufenden Monitorings. Andere Monitoringdaten sind nicht Voraussetzung, wohl aber Kontaktdaten zu den Absolventen und ggf. zu Arbeitgebern. Ein Vergleich mit Trainingsergebnissen bzw. Zwischen- oder *end-of-training*-Evaluationen während der Ausbildung kann sinnvoll sein.

Instrumente sind ein schriftlicher Fragebogen mit offenen und geschlossenen Fragen (ggf. auch online), Gruppeninterviews und Einzelinterviews. Die unterschiedlichen Perspektiven der Absolventen werden zusammengeführt und mit den Perspektiven der Beschäftigten, des Führungspersonals der Bildungseinrichtung/Beschäftigungsfördermaßnahme und evtl. Arbeitgebern verglichen. Weitere externe Perspektiven sind nicht Teil dieses Verfahrens.

d) Datenauswertung

Für die Datenauswertung werden Datenanalyseinstrumente verwendet, wie sie für standardisierte und offene Fragen üblich sind. Es gibt kein festgelegtes Raster. Unterschiedliche Perspektiven und Wahrnehmungen werden dargestellt und analysiert, um daraus Schlüsse zu ziehen. Validierung ist bei qualitativen Erhebungen möglich über Fokusgruppen mit Absolventen sowie Arbeitgebern. Es gibt i. d. R. keine systematische Überprüfung der Daten, aber Vergleiche der Einschätzungen von Absolventen mit Einschätzungen der Ausbilder zum Wert und zur Nutzung der Ausbildung. Systematische Überprüfungen sind bei höherem Aufwand möglich.

5. Nutzung der Ergebnisse

Das Instrument bringt hohe Transparenz und Nachvollziehbarkeit mit sich, sowohl bei externer wie bei Selbstevaluierung, da es auf Aussagen von direkten Nutzern basiert. Quantifizierbare und aggregierbare Daten sollten im Rahmen des Fragebogens erhoben werden. Die Verständlichkeit des Ergebnisses hängt davon ab, wie die Daten aufbereitet werden. Verfahren und Ergebnisse sind bei Selbstevaluierung darauf angelegt, dass sie von der Institution für die Weiterentwicklung der Programme und insbesondere der Curricula genutzt werden.

Nutzer sind zunächst nur die Durchführenden bzw. Auftraggeber, bei Selbstevaluierung also die durchführende Bildungsinstitution. Verbleibsstudien können für die Rechenschaftslegung gegenüber Gebern genutzt werden, dies kann aber die Objektivität beeinflussen.

6. Anforderungen

Der Aufwand für die Durchführung ist abhängig von der Zahl der Absolventen und weiteren Bedingungen. Bei Erstanwendung in einer Institution ist er höher als bei regelmäßiger Wiederholung. Bei schwerer Erreichbarkeit der Absolventen (hohe Mobilität) ist er höher als bei guter Erreichbarkeit. Zeit und Kosten sind stark von dem Vorhandensein von Kontaktdaten sowie Kommunikationsmöglichkeiten (z. B. Mobiltelefon) abhängig.

Tracer Studies setzen eine gewisse Kompetenz in der Sammlung und Analyse von Daten voraus. Mehrere Abteilungen in einer Institution müssen zusammenarbeiten.

7. Besondere Eignung und Grenzen des Verfahrens

Das Verfahren ist geeignet, die Wirkungs- und Kontextorientierung einer Berufsbildungseinrichtung bzw. eines Beschäftigungsförderprogrammes zu erhöhen und sie damit relevanter für ihre Umwelt zu machen. Es ist geeignet, Problemsituationen (z. B. eingeschränkter Zugang zu Beschäftigungs- und Weiterbildungsmöglichkeiten aufgrund von Gender und ethnischer, sozialer oder religiöser Zugehörigkeit) aufzuzeigen.

Literaturhinweis

Lange, R. (2001): Tracer Studies. Working instrument for a new orientation of vocational training. Stuttgart: AGKED/FAKT (www.fakt-consult.de).

Wirkungsorientiertes Monitoring (WoM) der GTZ

ARNO SCKEYDE

1. Allgemeine Informationen

Monitoring wurde bereits Mitte der achtziger Jahre als Instrument des Projektmanagement in die Arbeit der GTZ eingeführt. Die Idee war, von Anfang an Informationen für die zielorientierte Steuerung von Entwicklungsmaßnahmen und für die Berichterstattung bereitzustellen.

Nachdem Qualitätssicherung in den neunziger Jahren hauptsächlich auf *quality-at-entry* ausgerichtet war, trat einige Jahre später mehr die *quality-at-exit* in den Vordergrund. Mit der Einführung des neuen Auftragsrahmens AURA im August 2002 wurde die Wirkungsorientierung im Auftragsverfahren zwischen BMZ und GTZ maßgeblich. Das Erreichen der vorgegebenen entwicklungspolitischen Ziele rückte in den Mittelpunkt. Statt an erbrachten Leistungen wird der Erfolg der Maßnahmen nun an erreichten Wirkungen gemessen.

Wirkungsorientiertes Monitoring (WoM) ist kein grundlegend neues Instrument, sondern bezeichnet lediglich eine Form des Monitoring, das an Wirkungen orientiert ist – d. h. es schaut nicht nur auf die am Ende erzielte Wirkung (das wäre Wirkungsmonitoring), sondern besonders auch auf den Weg dahin, die Prozesse und Zwischenschritte. WoM dient der Steuerung, Rechenschaftslegung, dem Lernen und als Grundlage für Evaluierung von Entwicklungsmaßnahmen. Viele der international gebräuchlichen wirkungsorientierten Monitoringsysteme empfehlen eine bestimmte Schrittfolge, deren Anzahl Schritte variiert und zwischen denen es auch Ähnlichkeiten gibt. Das Verfahren in der vorliegenden Form wurde auf Grundlage der beschriebenen Anforderungen und der gemachten Erfahrungen in den Programmen von der GTZ entwickelt. Es ist immanenter Teil des Projekt- und Auftragsmanagements.

2. Anwendungsbereiche

Als Instrument der Steuerung, Rechenschaftslegung, des Lernens und des Wissensmanagements ist Wirkungsmonitoring in einer Vielzahl von Maßnahmen und Sektoren anwendbar. Die Frage nach dem Anwendungsbereich hängt nicht zuletzt von den schon vorhandenen Strukturen (z. B. des Partners) ab: Gibt es schon (Primär-)Daten oder Indikatoren auf Partnerebene, vorheriger Projekte oder muss eine *baseline* erhoben werden, von der ausgehend dann die Ziele definiert und Indikatoren zugeordnet werden?

Die GTZ verfolgt bei der Durchführung von Entwicklungsmaßnahmen grundsätzlich einen „Mehrebenenansatz“, d. h. – je nach Auftragsgestaltung – Interventionen auf der Makro-, Meso und ggf. auch der Mikroebene. Das beschriebene Vorgehen bei WoM ist ebenenübergreifend anwendbar.

Erfasst werden dabei sowohl Zustände (im Sinne von Wirkungen von Leistungen) als auch Prozesse und Prozessschritte, abhängig von der Frequenz, mit der Daten erhoben und ausgewertet werden.

3. Analyseansatz

Abhängig von der Datenlage und den geplanten Interventionsebenen der Maßnahmen werden gemeinsam mit den Partnern die Ziele und Indikatoren, die gemessen werden sollen, festgelegt.

Es wird unterschieden zwischen Wirkungen bis zur „Zuordnungslücke“ (direkte Wirkungen), die relativ eindeutig zuzuordnen sind, sowie solchen, die dahinter liegen (indirekte Wirkungen), wo eine plausible aber keine kausale Zuordnung mehr möglich ist (siehe Abbildung). Wichtig dabei festzuhalten ist: Für die GTZ steht das Monitoring der Wirkungen ihrer Leistungen im Vordergrund, nicht das Monitoring der Erbringung von Aktivitäten. Die (operative) Planung einer Maßnahme ist somit letztlich die Suche nach dem besten Weg, wie sich (eine) Hypothese(n) tatsächlich einstellt/-e(n). Ohne eine Überprüfung dieser Annahmen ist keine wirkungsorientierte Implementierung möglich.

Systemisch gesprochen beobachtet WoM aus zwei Perspektiven:

(1) Inwiefern trägt das Vorhaben zu Veränderungen im Umfeld bei?

Und:

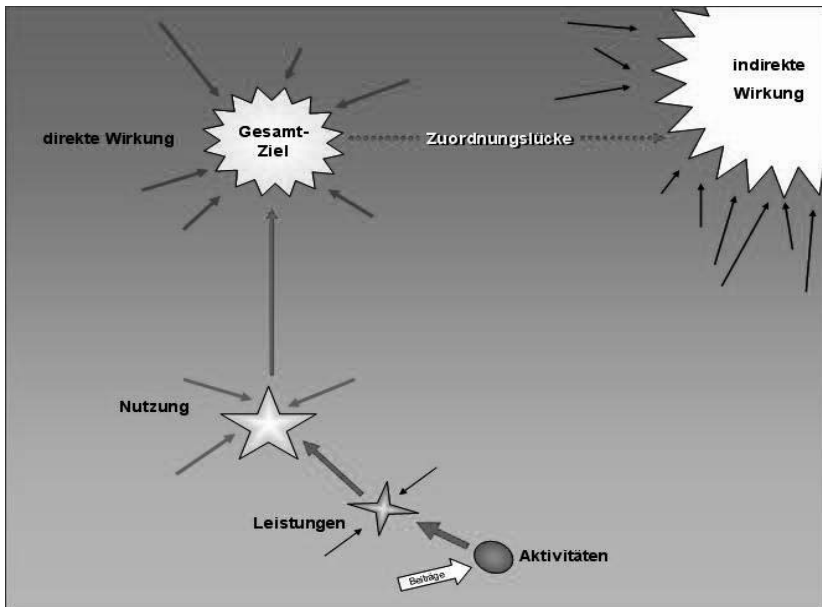
(2) Wie wirkt das Umfeld auf die Zielerreichung des Vorhabens ein?

Unter dem Begriff Wirkungen werden dabei Veränderungen gefasst, die kausal (direkt) oder plausibel (indirekt) auf die Intervention eines Vorhabens zurückgeführt werden können. Je höher aggregiert dabei die Wirkungsebene ist, desto schwieriger ist deren Zuordnung zu den Leistungen

der Entwicklungsmaßnahme. Viele positive und negative, unerwartete und unbeabsichtigte Wirkungen haben Einfluss auf die Wirkungskette. Im Umkehrschluss nimmt mit der Höhe der Aggregation der Wirkungen auch deren Beeinflussbarkeit ab. Dies stellt die GTZ vor zwei Herausforderungen:

- (1) Der kausale Nachweis der Zuordnung der Wirkungen zu den Leistungen des Programms muss erbracht werden (Zielebene) und
- (2) dort wo die Beeinflussbarkeit der Wirkungen hoch ist (niedrig aggregiertes Wirkungsniveau), ist es besonders wichtig, die Wirkungshypothesen zu überprüfen (Prozesssteuerung).

ABBILDUNG 17: Wirkungsmodell der GTZ



Gegebenenfalls müssen die angebotenen Leistungen angepasst werden. Des Weiteren steigt mit höherem Aggregationsniveau die Notwendigkeit der Beobachtung des (politischen und sozialen) Umfeldes und die Frage nach dessen Einfluss auf die Wirkungen.

Zentrale Frage für das Management der Maßnahme ist dabei immer, ob sich diese noch im Zielkorridor bewegt, ob sich die Rahmenbedingungen verändert haben und ob im Verlauf der Durchführung Wirkungen auftreten,

die die Zielerreichung gefährden. Gegenstand der Analyse sind dabei nicht nur die erreichten, sondern auch unerwartete und negative Wirkungen, ebenso wie Leistungen Aktivitäten, Finanzierungen und Risiken. WoM folgt dabei dem Prinzip, mit möglichst wenig Input und möglichst breiter Nutzung vorhandener Ressourcen möglichst verlässliche Ergebnisse zu generieren.

4. Durchführung

Der Aufbau eines WoM in einem Projekt bedeutet in der Regel nicht die Einführung eines zusätzlichen Instrumentes, sondern ist vielmehr Neuausrichtung der bereits vorhandenen Monitoringaktivitäten auf Wirkungen, d. h. der Inwertsetzung von Partner-/Monitoringsystemen: beobachtet wird nicht primär was gemacht wurde, sondern welche Veränderungen bewirkt werden. Kernelemente dieser Betrachtung sind:

- Zielsetzung und genauer Zweck des Monitoring;
- die konkreten Inhalte;
- eine entsprechende Struktur, die die Inhalte systematisiert;
- ein Verfahren, in welchem die regelmäßige Beobachtung und Auswertung organisiert ist.

a) Verfahrensablauf

WoM lässt sich in 7 Schritten beschreiben (siehe Abbildung):

1. Wirkungsketten und Systemgrenzen als Grundlage des Monitoring

Mit den Partnern erarbeitet, machen Wirkungsketten und Systemgrenzen deutlich, für welche Ziele und Umsetzungsprozesse eine gemeinsame Verantwortung übernommen wird. Leitfragen:

- Wer gehört zum Vorhaben? Wo beginnt sein Umfeld?
- Klärung der Beteiligten, der Einflussfaktoren, des erwarteten Nutzens, usw.

2. Interessen, Erwartungen und Beiträge zum Monitoring

Die Klärung von Interessen und Erwartungen an das gemeinsame Monitoring mit den Partnern und weiteren Beteiligten zielt darauf ab, das Monitoringsystem auf die Ziele und Bedarfe seiner potenziellen Nutzer (also langfristig der Partnerinstitutionen) auszurichten und seine Akzeptanz innerhalb und außerhalb der TZ-Maßnahme zu erhöhen. Leitfragen:

- Wer braucht welche Informationen zur Steuerung, zur Rechenschaftslegung oder für andere Zwecke?
- Welche Interessen und Erwartungen haben potenziell Beteiligte am Monitoring der Wirkungen?

3. Beobachtungsfelder des Monitoring

Für die Steuerung einer Entwicklungsmaßnahme und die wirkungsorientierte Berichterstattung ist es nicht nötig, alle von ihr ausgelösten Veränderungen zu verfolgen. Projekte und Programme müssen jedoch jederzeit in der Lage sein, Auftraggebern, Partnerorganisationen und weiteren Gebern Auskunft darüber zu geben, ob die vereinbarten Ziele erreicht werden, und welche Wirkungen bereits erzielt wurden. Dies geschieht anhand der Erfüllung der definierten Indikatoren. Leitfragen:

- Welche Wirkungshypothesen sind für die Zielerreichung besonders wichtig?
- Wo sind Wirkungsketten miteinander verknüpft? Welche Veränderungsschritte können nur durch zwei oder mehrere zusammenlaufende Veränderungsprozesse erreicht werden?

4. Zielindikatoren, Prozessindikatoren und Monitoringfragen

Um zu erkennen, ob eine Entwicklungsmaßnahme die angestrebten Ziele erreicht, sind Indikatoren nötig. Leitfragen:

- Was soll sich ändern?
- Bei wem soll sich die Änderung auswirken?

5. Struktur und Prozesse des Monitoring

Nach der Klärung der Rollen bei Datensammlung, -verarbeitung und Interpretation der Daten hinsichtlich spezifischer Indikatoren und Fragestellungen müssen die Verantwortlichkeiten für Elemente des Monitoringsystems festgelegt und übertragen werden. Aspekte:

- Planung und Koordination;
- Design und Unterstützung einzelner Monitoringaktivitäten;
- Monitoring und Qualitätskontrolle der Monitoringaktivitäten;
- Berichterstattung;
- Dokumentation und Informationsvermittlung.

6. Datenerhebung

Die Datenerhebung wird entsprechend der operativen Planung der Monitoringaktivitäten durchgeführt und anhand des Operationsplans und regelmäßiger Teambesprechungen gesteuert.

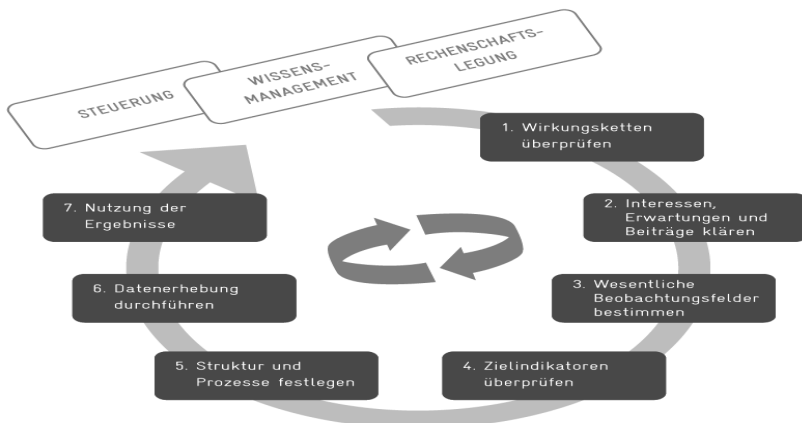
- Welche Daten werden für die vom Auftraggeber bzw. dem Aufsichtsgremium der TZ-Maßnahme vorgegebenen Indikatoren benötigt?
- Welcher Grad an Genauigkeit, Zuverlässigkeit und Repräsentativität wird angestrebt?

7. Nutzung der Monitoringergebnisse

Der Nutzen des Monitoring stellt sich erst mit der Nutzung seiner Monitoringergebnisse ein. Entscheidend ist, dass es gelingt, ein Klima des Vertrauens im Team zu schaffen sowie die Voraussetzungen für einen konstruktiven Dialog. Das sind die Grundlagen, um Informationen des Monitoringsystems für

- die Steuerung der Entwicklungsmaßnahme,
- die Rechenschaftslegung gegenüber Auftraggebern, Steuerungsgremien, anderen Gebern und Öffentlichkeit im Partnerland und in Deutschland und eventuell für
- das interne wie externe Wissensmanagement zu nutzen.

ABBILDUNG 18: 7 Schritte eines wirkungsorientierten Monitoring in der GTZ



b) Akteure

Der Partner, soweit er in die Durchführung der Entwicklungsmaßnahme involviert ist, sollte bei der Entwicklung und Umsetzung eines WoM-Systems in einem Vorhaben von Anfang an eine zentrale Rolle spielen, insbesondere wenn es möglich ist, auf dessen vorhandenen Analyseinstrumente, Monitoringelemente und vorliegende Daten zurückzugreifen und die eigenen Monitoringaktivitäten dort zu integrieren. Auch in Bezug auf *alignment* (Pariser Deklaration) steht die Frage, wie die Maßnahme zu den nationalen Zielen der Partner beiträgt, im Fokus. WoM kann dazu Hinweise zur Plausibilität liefern.

c) und d) Datenerhebung und Datenauswertung

In welcher Form, auf welchem Weg und mit welcher Frequenz die Daten erhoben und ausgewertet werden, hängt in erster Linie von den Indikatoren ab, die (möglichst in Kooperation mit dem Partner) entwickelt wurden.

Wenn keine Ausgangsdaten vorliegen, sollte in jedem Fall eine *baseline* erhoben werden.

5. Nutzung der Ergebnisse

Die Nutzung der Ergebnisse von WoM dienen der Steuerung, der Rechenschaftslegung, dem Lernen und dem Wissensmanagement (siehe Schritt 7 des Verfahrens). Die dezentralen und z. T. sehr regional- und sektorspezifischen Erfahrungen werden als *good practice* nach einem festgelegten Format aufbereitet (*lessons learned*) und zentral zur Orientierung zur Verfügung gestellt.

6. Anforderungen

Formale Voraussetzung für WoM ist natürlich die Bereitschaft aller Beteiligten, WoM ernsthaft (und ehrlich) zu betreiben. Bei der Weltbank ist daher der erste Schritt bei WoM die Prüfung der *readiness*. In der GTZ ist es für jedes Vorhaben verbindlich, WoM durchzuführen. Dies wird als eine für das Projekt existentiell wichtige Aufgabe und daher als Führungsaufgabe gesehen. Die Anforderungen hängen vom Kontext ab; je nach Ausrichtung und Ausstattung der Maßnahme sind 5-15% des operativen Budgets einer Maßnahme für Monitoring vorzusehen.

7. Besondere Eignung und Grenzen des Verfahrens

Das beschriebene Wirkungsorientierte Monitoring bezieht sich in erster Linie auf bilaterale EZ-Maßnahmen, kann aber auch im Rahmen von Multi-Geber-Ansätzen verwendet werden. Allerdings sind hier die Qualität und Verbindlichkeit der Abstimmung der beteiligten Partner entscheidend.

Ein generelles Problem ist die Erfassung der indirekten und hoch aggregierte Wirkungen. Dies ist mit dem Anspruch, Kausalzusammenhänge darzustellen, prinzipiell schwierig oder – im Rahmen der heutigen (i. d. R. noch stark fragmentierten) Akteurslandschaft – schwierig oder unmöglich. Hier können Verfahren der Ex-ante-Wirkungsprognose (wie PSIA) hilfreich sein.

Literaturhinweis

Gesellschaft für Technische Zusammenarbeit (GTZ) (2008): Wirkungsorientiertes Monitoring. Leitfaden für Vorhaben der Technischen Zusammenarbeit. Eschborn.

2.3 Weitere Ansätze der Wirkungsanalyse

Im Folgenden werden einige weitere Ansätze der Wirkungsanalyse für bestimmte Zwecke und Sektoren vorgestellt, die sich nicht strikt entlang der Struktur der Steckbriefe (siehe Abschnitt 2.1) darstellen lassen und/oder die sich derzeit noch im Erprobungsstadium befinden.

Wirkungsanalysen von EZ in Konfliktsituationen

MICHAELA ZINTL

In einer Vielzahl von Partnerländern des BMZ herrscht latenter oder akuter Konflikt. Der Untersuchung von Wirkungen von Entwicklungszusammenarbeit in Konfliktsituationen kommt daher besondere Bedeutung zu. Einen neuartigen Ansatz, die (kumulativen) Wirkungen von Entwicklungszusammenarbeit in Konfliktsituationen zu untersuchen, stellt die im Herbst 2006 begonnene Wirkungsanalyse in Nordost-Afghanistan dar. Sie wird in Kooperation mit einem Forschungsteam der FU Berlin unter Leitung von Christoph Zürcher durchgeführt. Sie soll relevante Steuerungsinformationen liefern und darüber hinaus der Methodenentwicklung (Entwicklung eines reproduzierbaren Verfahrens zur Wirkungsanalyse im Konfliktkontext) dienen. Durch die Beteiligung und Schulung einer lokalen Forschungseinrichtung wird überdies ein Beitrag zum *capacity development* geleistet.

Entsprechend des zugrundeliegenden Wirkungsmodells soll, ausgehend von einer bereits existierenden Konfliktanalyse und den erklärten Zielen der Gebergemeinschaft, der Einfluss von Entwicklungszusammenarbeit getestet werden im Hinblick auf: a) Einstellungen der betroffenen Bevölkerung gegenüber westlichen Akteuren; b) Einstellungen gegenüber fremden Truppen; c) Zufriedenheit mit der Arbeit von Entwicklungsorganisationen in zentralen Feldern wie Wasser, Elektrizität, Bildung, Gesundheit; d) Einschätzung von Sicherheit; e) kommunale Organisationsfähigkeit; und e) die Fähigkeit der Regierung, Basisversorgung sicher zu stellen.

Da keine Daten zur Beantwortung dieser Forschungsfragen existierten, wurde ein Survey anhand eines standardisierten Fragebogens, der zuvor getestet wurde, unter 2.000 Haushalten in 80 Dörfern durchgeführt. Es wurde sichergestellt, dass die Stichprobe für die Gemeinden repräsentativ ist, so dass Vergleiche auf Gemeindeebene möglich sind.

Erfasst wurde weiterhin die Anzahl und Art der in den Gemeinden durchgeführten Entwicklungsprojekte sowie die Perzeption der Nützlichkeit dieser Projekte durch die Bevölkerung. Zusätzlich wurden für alle 80 Dörfer Profile erstellt (Größe, ethnische Zusammensetzung, Konflikthistorie, Entwicklungsprojekte usw.). Ergänzend wurden vorhandene Daten (z. B. Vulnerabilitätsstudien) hinzugezogen. Interviews mittels Fokusgruppen und mit Schlüsselpersonen ergänzen die standardisierte Befragung.

Die Ergebnisse werden in einem Datensatz zusammengefasst und mittels deskriptiver und induktiver Statistik ausgewertet. Für die Messung der Wirkungen der Entwicklungszusammenarbeit – anhand von Perzeptionen der lokalen Bevölkerung – wurden OLS-Regressionen berechnet und externe Faktoren kontrolliert (geographische Lage, Vulnerabilität der Haushalte, Ethnizität usw.).

Wie bei anderen Wirkungsanalysen auch, ist die Zuordnung beobachteter Veränderungen eine Herausforderung. So wurde geprüft, ob die Bereitstellung von Hilfe von Merkmalen der Zielbevölkerung beeinflusst ist (*selection bias*). Da in der untersuchten Region alle Dörfer irgendeine Form der Hilfe erhalten haben, kann keine Vergleichsgruppe „ohne Hilfe“ gebildet werden, allerdings können unterschiedliche Formen der Unterstützungsleistung variiert werden. Entsprechend wird 2009 ein zweiter Survey in denselben Gemeinden durchgeführt, ergänzt durch qualitative Fallstudien. Vierteljährliche *surveys* zu externen Ereignissen werden während der gesamten Laufzeit durchgeführt.

Evaluierungen in Konfliktzonen sind in jeder Hinsicht herausfordernd, was sich unter „keine Straßen, keine Namen, keine Daten“ zusammenfassen lässt.

Die Ergebnisse der ersten Untersuchungsrunde und die methodischen Grundlagen sind in einem Zwischenbericht und einem ersten Bericht zur Methodik zusammengefasst (Zürcher, Köhler und Böhnke (2007), Zürcher und Köhler (2007). Der Endbericht und aktualisierte Methodenbericht werden bis Ende 2009 vorliegen.

Wirkungsanalyse von Ernährungssicherungsmaßnahmen

CHRISTIAN BERG

Für Vorhaben der Ernährungssicherung außerhalb der humanitären Soforthilfe, d. h. im Rahmen der entwicklungsorientierten Not- und Übergangshilfe (EON, ENÜH), existieren keine spezifischen Wirkungsanalyseverfahren. Allerdings gibt es Instrumente zur ex-ante-Erfassung der Ernährungslage in einer Zielbevölkerung und -region, die sich prinzipiell auch für eine Beurteilung der Wirksamkeit der jeweiligen Maßnahmen eignen. Voraussetzung dafür wäre eine Wiederholung der Erhebungen zu späteren Zeitpunkten, verbunden mit der Erfassung und Gewichtung externer Einflussfaktoren auf Änderungen in der Ernährungssituation.

Rapid Nutrition Assessment (RNA) oder auch **Rapid Assessment of Nutrition (RAN)** ermöglicht einen schnellen – quantitativen wie qualitativen – Überblick über die Ernährungslage gefährdeter Gebiete und betroffener Bevölkerungsgruppen (siehe <http://www.methodfinder.net/index.php?page=methods&methodID=52&cat=briefdescription>; <http://www.dppc.gov.et/pages/ENCU.htm>). Das Verfahren liefert auf der Basis von Sekundärdatenanalyse, Schlüsselinformanten-Interviews, Zielgruppenbefragungen, direkten Beobachtungen und Messungen anthropometrischer Indikatoren direkte Informationen zur akuten Notlage in einer Region, aus denen angemessene Entscheidungen über Interventionen abgeleitet werden. RNA/RAN ist kein Ersatz für detaillierte Ernährungsstudien und die Ergebnisse der Erhebung können durch nicht-repräsentative Auswahl von Teilnehmern/-innen verzerrt sein.

Household Food Economy Survey (HFES) oder auch **Household Economy Approach (HEA)** liefert detaillierte Daten zur Nahrungsmittelverfügbarkeit, Zugang zu und Verwertbarkeit von Nahrungsmitteln, Einkommens- und Ausgabenstruktur und der Vulnerabilität verschiedener Wohlstandsgruppen einer Bevölkerung (siehe z. B. http://www.savethechildren.org.uk/en/54_4200.htm). Das Verfahren kann einen Beitrag zur Verbesserung der Interventionen und Strategien im Bereich Nahrungsmittelsicherheit sowie zur Bedarfsanalyse der weiteren Projektplanung und Prioritätenbestimmung der Projektziele leisten. Die Erhebung wird von einer qualifizierten und erfahrenen – in der Regel einheimischen – Fachkraft geleitet und hat als wesentliche Elemente:

- Eine Einteilung des Untersuchungsgebietes in Food Economy Zones (FEZ): Dabei werden sozio-ökonomische und agro-ökologische Fak-

toren zugrunde gelegt. Das Untersuchungsgebiet wird in Zonen eingeteilt, die in sich möglichst homogen sind, beispielsweise in eine Hochland-Zone und eine Zone der mittleren Höhen.

- Definition von Wohlstandskategorien: In Zusammenarbeit mit den Zielgruppen (anhand von Gruppendiskussion) werden die Kategorien reich – mittel – arm mit jeweils korrespondierenden Kriterien definiert, beispielsweise Besitz von Zugochsen, Land.
- Auswahl der Stichprobe: Im Unterschied zu repräsentativen Haushaltsuntersuchungen werden relativ wenige Personen – als Schlüsselpersonen – befragt. Diese werden so ausgewählt, dass alle Wohlstandskategorien vertreten sind (kleine Stichprobe, 10-20 Befragungen).
- Durchführung der Interviews: Bei der Befragung der ausgewählten Personen werden verschiedene PRA-Instrumente angewandt. Es werden die Bereiche Nahrungsmittelverfügbarkeit, Einkommensquellen und Ausgabenstruktur untersucht.

HFES ist in quasi-kontinuierlichen Nothilfesituationen (*protracted relief and recovery*) sehr hilfreich, da Einkommen, Auskommen und Bewältigungsstrategien (*coping strategies*) von Haushalten genau analysiert werden. Allerdings ist die Vorab-Definition eines Haushalts (wer gehört dazu?) und die Interpretation der zumeist qualitativen Daten in der Regel schwierig.

Erfassung der Wirkungen von (Multi-)Sektorprogrammen

SUSANNE NEUBERT

Die Schweizer Entwicklungsbehörde (DEZA) hat 2007 ein Instrument entwickelt, das explizit zur Informationsgewinnung über die Einbettung der Programme in einen größeren Kontext dient (Orientierung auf die Paris-Erklärung, Kohärenz der Programme mit größeren international akzeptierten Konzepten) und auf die Erhebung der Wirkungen aller Maßnahmen in einem Sektor (**Rapid Sectoral Assessment, RSA**) bzw. des gesamten Portfolios in einem Land abzielt (**Rapid Portfolio Assessment, RPA**).

Das Verfahren beruht auf schriftlichen Einzelbefragungen (mit Hilfe von Excel[®]-sheets) und ist primär von Programmverantwortlichen und ent-

sprechenden Partnern zu beantworten. Das kompetente Ausfüllen der Bögen setzt einen Überblick über den gesamten Sektor bzw. das Portfolio des Geberlandes und der Politiken des Partnerlandes voraus. Daher kommen nur sehr wenige Personen zur Befragung in Frage. Dem Zuordnungsproblem wird begegnet, indem jeweils die Gesamtwirkung von Maßnahmen und die Einzelwirkungen der Maßnahmen eines Geberlandes getrennt abgefragt werden. Zudem werden explizit Begründungen oder Belege für Wirkungsaussagen erfragt. Qualitative Aussagen werden mit Punktwertungen kombiniert. Als Ergebnis wird die Entwicklung der Indikatoren tabellarisch oder/und grafisch in Form von Säulendiagrammen aufgezeigt. Eine explizite Darstellung der zeitlichen Dimension entgegen einer Zielagenda wird nicht vorgenommen.

2007 wurde RSA vom DIE modifiziert und im Rahmen einer Analyse der Wirksamkeit der EZ im Wassersektor der DEZA/Seco eingesetzt. Die Modifizierung bestand in erster Linie in der stärkeren Betonung qualitativer Aspekte und in der Integration stärkerer Begründungszwänge, damit Wirkungen nicht nur postuliert, sondern auch fundiert werden. RSA ist geeignet, als übergeordnetes „Dach“ für Sektor- und Portfolioanalysen zu dienen, die durch systematische Befragungen auf der Bevölkerungsebene ergänzt werden können / sollten. RSA eignet sich daher als Bestandteil eines Methodenmix.

Der Gewinn von RSA liegt darin, dass nicht nur begründete Wirkungsvermutungen vermittelt, sondern v. a. Auskunft zur strategischen und nationalen Einbettung der Maßnahmen gegeben wird. Damit hat das Verfahren einen Alleinstellungscharakter. Die grafische Darstellung ermöglicht eine schnelle Erfassung der Bewertungen. Der verbundene Aufwand ist gering und daher ist das Verfahren kostengünstig.

Allerdings ist von einem positiven Bias bei den Antworten auszugehen. RSA beinhaltet kein Konzept zum Umgang mit unterschiedlichen Angaben / Kontroversen und enthält kein explizites System zur Schlussbewertung der Ergebnisse.

Multi-Sectoral(MS)-Monitoring wurde multilateral unter Beteiligung der DEZA im Rahmen ihrer Mitarbeit an multilateralen Programmen im Jahr 2007 entwickelt. Es geht erstens um die Erhebung der Gesamtwirkungen dieser Programme und zweitens um den Politikeinfluss eines spezifischen Mitglieds / eines Geberlandes im Hinblick auf die Ausgestaltung der Programme. Der zweite Aspekt wurde bei einer Modifizierung der Methode durch das DIE hinzugefügt. Zu diesem Zweck wurde das Zielkonzept „Politikeinfluss“ zunächst operationalisiert und Kategorien für „Politikeinfluss“ entwickelt. Die Unterscheidungskriterien für Politikeinfluss, wie sie im entsprechenden Abschnitt weiter unten dargelegt sind, wurden mit Beispielen /

Belegen gefüllt: „Beeinflussung des Diskurses, der Moden der Institutionen, der Normen, bindender Regelwerke, bestimmter Politikinstrumente und schließlich der Praxis auf Bevölkerungsebene“.

Die Methode basiert auf semi-strukturierten Telefoninterviews mit Stakeholdern des zu evaluierenden Programmmitglieds sowie auf ergänzenden Interviews (kontrastierenden Vergleichen) mit anderen, explizit auch kritischen Mitgliedern (anderer Geberländer) und mit kooperierenden Vertretern der Partnerländer. In Anbetracht der immer knappen Zeit der Telefonpartner konzentriert sich die Befragung auf die Nennung der drei wichtigsten positiven und negativen Wirkungen (*highlights* und *setbacks*) sowie auf die Bewertung des spezifischen Beitrags des zu evaluierenden Mitglieds unter Nennung der Möglichkeiten, in welcher Form Politikeinfluss geübt werden kann. Die Nennung jeweils konkreter Beispiele ist hierbei besonders wichtig.

MS-Monitoring besitzt ein Alleinstellungsmerkmal, weil hier die Wirkungsanalyse multilateraler Programme im Vordergrund steht. Das Verfahren hat sich als praktikabel, einfach und günstig erwiesen (Telefoninterviews und Dokumentenanalyse).

Das Verfahren wurde zwar im Rahmen der Datenerhebung für den Wirkungsbericht der DEZA/Seco erprobt, kann aber formal noch besser beschrieben und standardisiert werden. Bisher können Wirkungen noch nicht klar genug von Outputs getrennt werden. Spezifische Beiträge sind oft mit Gesamtwirkungen (exogenen Faktoren) verwoben. Die Ergebnisse können einen positiven Bias haben, der nur auszuschließen ist, wenn eine relevante Anzahl von Interviews auch mit unabhängigen Partner/-innen geführt wird.

Methodisches Vorgehen bei der Ermittlung von Politikeinfluss

SUSANNE NEUBERT

Entwicklungszusammenarbeit besteht heute immer häufiger aus Vorhaben der Politikberatung und wird zunehmend im Rahmen von Netzwerken und mit mehreren Gebern durchgeführt. Dabei geht es nicht in erster Linie um die Durchführung praktischer Interventionen, sondern um die Beeinflussung der Partnerregierung oder eines Netzwerkes zur Institutionalisierung und

Umsetzung spezifischer Entscheidungen oder Reformen auf nationaler bis lokaler Ebene.

Möchte man daher die Wirksamkeit von Vorhaben der Politikberatung messen, muss man zunächst den Grad der Politikbeeinflussung auf die gemeinsame Entscheidungsfindung operationalisieren.

Es ist hier aus methodischer Perspektive gleichgültig, ob es sich um Strategien multilateraler Netzwerke oder um Regierungspolitiken und -programme handelt. Im ersten Fall wird der Einfluss einzelner Netzwerkmitglieder auf Entscheidungen und Maßnahmen dieser Netzwerke untersucht, im zweiten Fall wird der Einfluss von Gebern oder/und Beratern auf das Regierungshandeln in den Partnerländern analysiert. Beide Maßnahmentypen der EZ können aus finanziellen, personellen oder beratenden Leistungen oder einer Kombination daraus bestehen.

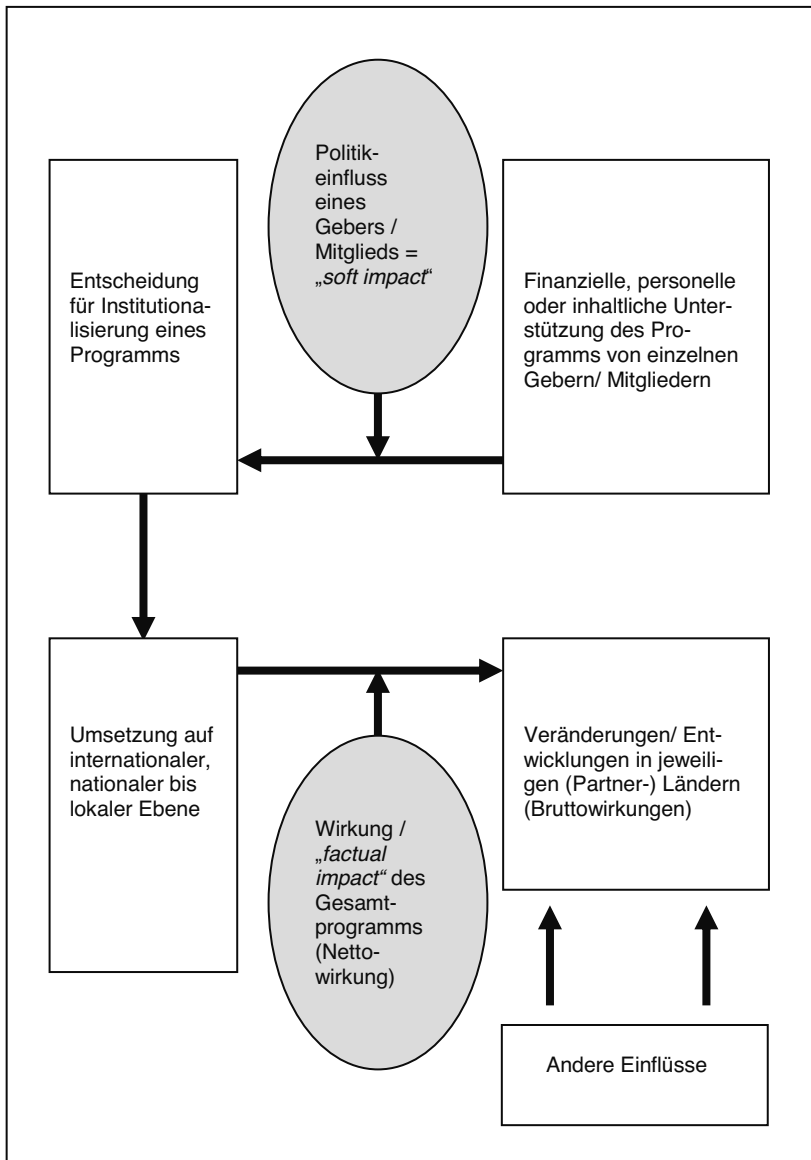
Konzeptionalisierung

Bei beiden Maßnahmentypen werden die folgenden Wirkungsziele verfolgt:

1. Der spezifische Beitrag des Akteurs soll im Rahmen der internationalen Politik oder der Politik des Partnerlandes sichtbar werden, d. h. die Strategien oder Entscheidungen auf die man sich einigt, sollen mehr oder minder „seine Handschrift“ tragen. Denn nur wenn der spezifische Beitrag des Akteurs „einen identifizierbaren Unterschied macht“, kann der Ressourceneinsatz legitimiert werden.
2. Außerdem soll der Ressourceneinsatz des Mitglieds oder Gebers möglichst wirksam sein. Die verfochtenen Politikansätze und unterstützten Maßnahmen müssen daher im Partnerland oder auf internationaler Ebene tatsächlich institutionalisiert und umgesetzt werden. Auch sollten sie, und dies ist der eigentliche Sinn, auch *on the ground*, d. h. auf Bevölkerungsebene wirksam werden, und nur so sind sie letztlich auch zu rechtfertigen.

Während das zweite Ziel bei allen Vorhaben der EZ zu berücksichtigen ist und mit bestehenden Verfahren der Wirkungsanalyse, wie sie in diesem Handbuch beschrieben werden, evaluiert werden kann, beschreibt das erste Ziel eine neue Dimension, die die Feststellung des Einflusses des entsprechenden Akteurs verlangt, der von ihm in Richtung der dann tatsächlich gemeinsam getroffenen Entscheidungen ausgeübt wurde. Dies ist wichtig, denn wenn eine wirksame/ richtige Politik auch ohne externen Einfluss/Unterstützung durchgeführt worden wäre oder wenn ein Netzwerk gleichermaßen auch ohne dieses eine Mitglied funktionierte, dann wäre die Ressourcenallokation suboptimal zu bewerten.

ABBILDUNG 19: Relevante Wirkungsdimensionen erster und zweiter Ordnung bei Vorhaben mit zwei und mehreren Partnern



Neben der schieren Feststellung, ob ein Politikeinfluss des Gebers/Mitglieds existiert oder nicht, interessiert bei der Methodik die graduelle Abstufung dieses Einflusses, d. h. die Einflussstärke. Hierfür und wie sich Einfluss manifestiert, d. h. welche Evidenzen hierfür bestehen, gibt es keine etablierten Vorlagen. Daher wird hier ein neues Bewertungsraster vorgestellt, das einen ersten Schritt der Operationalisierung darstellt.

Demnach besteht der Minimalanspruch einer Einflussnahme darin, die eigenen Konzepte, Begriffe oder Themen in die Diskurse mit den Partnern oder Netzwerken einzubringen, d. h. zunächst darauf erfolgreich einzuwirken, dass diese darin Eingang finden (*agenda setting*). Schon höher erscheint allerdings der Einfluss, wenn durch diese Diskurse neue Moden entstehen, die sich langsam aber sicher in einen allgemeinen Trend entwickeln (*trend setting*). Noch stärker könnte man allerdings den Einfluss ansehen, wenn durch die Mitarbeit neue Normen gebildet werden, die als solche in die nationalen Politiken eingehen (Normbildung). Eine noch weitergehende Wirkung würde darüber hinaus erzielt, wenn sich die Institutionenlandschaft im Partnerland im Sinne der eingebrachten Position verändert, z. B. durch den Aufbau neuer Organisationen und/oder wenn bindende Abkommen/Konventionen in Zielrichtung zustande kämen (Beeinflussung bindender Regelwerke). Das höchste Maß der Beeinflussung wäre allerdings erreicht, wenn die gemeinsame Arbeit zur Umsetzung der neuen Politikinstrumente führte.

Man muss hierbei allerdings nicht meinen, dass das höchste Maß der Beeinflussung unbedingt per se das Ziel darstellt, sondern besser und realistischer ist es vielmehr, wenn beide Partner oder alle Mitglieder des Netzwerkes ungefähr gleichgewichtig an Entscheidungen beteiligt sind und dabei auch jeweils einen gewissen Einfluss ausüben konnten.

Zur Bewertung der Leistung eines Akteurs und dessen Politikeinfluss ist es zudem wichtig festzustellen, inwiefern eine Empfehlung / ein Einfluss eigentlich quer zum allgemeinen Denken stand oder ob es dieses allgemeine Denken selbst nur widerspiegelt. Allgemein bereits akzeptierte Vorstellungen durchzusetzen dürfte somit wesentlich weniger Einsatz erfordern als zum Umdenken anzuregen und Vorstellungen durchzusetzen, die bisher nicht en vogue waren.

ABBILDUNG 20: Raster zur Messung von Politikeinfluss

Minimum der Beeinflussung



Grad der Beeinflussung	Beispiel aus dem Wasserbereich
1. Beeinflussung des Diskurses a) Themensetzung wird beeinflusst oder/und qualitativ verbessert (Agendasetting, Wissensverbreitung, Intensivierung des Dialogs). b) Zusammensetzung und Zahl der beteiligten Dialogpartner werden beeinflusst (funktionstüchtiges, kompetentes Netzwerk/Entscheidungsgremium).	a) Das Thema „Menschenrecht auf Wasser“, das im besonderen Interesse des Akteurs ist, wird auf die Agenda des Netzwerkes/der Regierung gesetzt. b) In ein Politikforum über Wasser werden Teilnehmer der Zivilgesellschaft neu einbezogen. Es wird eine digitale Dialogplattform geschaffen, die öffentlich einsehbar ist.
2. Beeinflussung der Moden Spezifische Konzepte/ Begriffe etc. werden allgemein akzeptiert (Trendsetting).	Z. B. „Sanitation for all“ wird als internationales oder als nationales Entwicklungsziel in Konzepte übernommen (MDGs, PRSPs).
3. Beeinflussung der Institutionen(-landschaft) Neue Institutionen/ Netzwerke/ Organisationen werden gegründet.	Ein Sekretariat/eine Kommission wird etabliert (z. B. World Commission on Dams, WCD). <i>Oder:</i> Neue Institutionen im Sinne des IWRM werden gegründet (z. B. Flussgebietsorganisationen).
4. Beeinflussung der Normbildung Spezifische Regelwerke werden Bestandteil der Netzwerk- oder Regierungspolitik.	Empfehlungen der WCD werden in die nationale Sektorpolitik als Norm unverbindlich übernommen.
5. Beeinflussung bindender Regelwerke International: Bindende Abkommen werden mit begründet und/oder ihre Bekanntheit und Einhaltung vorangetrieben. National: Gesetzgebungen oder Reformen werden tatsächlich novelliert, angestoßen.	Empfehlungen der GTZ werden in die Wassersektorreform aufgenommen/übernommen.
6. Beeinflussung des Einsatzes bestimmter Politikinstrumente Angestrebte Politikinstrumente werden tatsächlich institutionalisiert.	Nachfragesteuerung des Wasserverbrauchs wird auf der institutionellen Ebene umgesetzt, z. B. werden Preise auf Rohwasser erhoben.
7. Beeinflussung der Umsetzung in die Praxis Politikinstrumente werden in der Praxis umgesetzt/enforcement.	Kommerzielle Wasserverbraucher messen und bezahlen die Wasserentnahmen.

Maximum der Beeinflussung

Es ist daher wichtig einen Orientierungsrahmen heranzuziehen, mit dessen Hilfe man feststellen kann, ob gewisse Entscheidungen eigentlich sinnvoll und gewünscht und inwiefern sie tatsächlich im Sinne des Gebers/Mitglieds und/oder Partners sind. Gleichmaßen ist es wichtig, die Ausgangslage zu verstehen und die Governance-Strukturen im jeweiligen Land zu kennen. Die Politikberatungsinhalte müssen sich daran orientieren (Kontextorientierung) und auch die Anspruchshaltung auf Veränderung.

Hinzu kommt die Berücksichtigung des Zeitrahmens: Institutionelle Veränderungen benötigen in der Regel wesentlich längere Zeit als die Durchführung von z. B. Infrastrukturmaßnahmen. Es sind daher realistische Zeitansprüche zu stellen, die je nach Maßnahme und Zielsetzung zwischen einem Jahr und mehreren Jahrzehnten liegen können.

Erhebung der Daten

Die nächste Frage ist nun, wie die jeweilige Einflussstärke zu messen ist, bzw. wie die entsprechenden Daten erhoben werden können und die Zuordnungslücke, die auch hier auftritt, überbrückt werden kann.

Während sich im Falle multilateraler Vorhaben Telefoninterviews mit mehreren beteiligten Netzwerkmitgliedern anbieten (Besuche vor Ort wären hier sehr aufwändig), müssen bei bilateralen Vorhaben die Perspektive des Akteurs/Beraters selbst sowie die der Regierung/des Partners erfasst werden. Diese können sowohl per Telefon als auch mit Hilfe persönlicher Interviews erhoben werden.

Um einen Bias oder eine Über- oder Unterschätzung des Einflusses eines Akteurs auf Entscheidungen zu vermeiden, sollten möglichst unterschiedliche Stakeholder mit verschiedenen Positionen oder Rollen in die Befragung einbezogen werden. Anschließend bieten sich kontrastierende Vergleiche der Perspektiven der Gesprächspartner an, die mit Hintergrunddokumenten oder Vor-Ort-Besuchen abgewogen werden.

Die Zuordnungslücke kann überbrückt werden, indem bei der Befragung zunächst die Veränderung der Politik erhoben wird und dann die besondere Rolle des zu evaluierenden Akteurs genau hinterfragt wird.

Für die Befragung wird zudem ein Orientierungsrahmen benötigt (s. o.), der das Zielkonzept umreißt. Solch einen Rahmen kann die jeweilige multilaterale Strategie des Gebers darstellen oder das Länder- bzw. Sektor-konzept.

Für das *cross-checking* können Dokumente des Akteurs hinzugezogen werden, die z. B. belegen, dass bestimmte Normen (z. B. das Menschenrecht auf Wasser) bereits im Vorfeld des Politikberatungsprojekts ein Anliegen des Akteurs waren. Allerdings sollen sich bei echten Diskursen Positionen des Akteurs auch verändern können. Hier sollte es also nicht darum

gehen, die vollständige Durchsetzung unverrückbarer Positionen anzustreben. Tritt eine solche Veränderung der Position auf, sollte dieser Prozess allerdings kenntlich und transparent gemacht werden und sich auch in den geplanten zukünftigen Dokumenten des Partners wiederfinden lassen.

Die so erhobenen Daten werden nun abgewogen und den Stufen für Politikeinfluss (dem Raster) zugeordnet. Kontroverse Ergebnisse werden als solche kenntlich gemacht.

Die Schlussbewertung orientiert sich daran, welcher Grad an Einfluss überhaupt erwartet werden konnte. Daher müssen bei der Befragung sowohl Daten zur Größe der Akteure (Budgets), als auch zu ihrem Image (zugeschriebene Kompetenz) und zu ihrer informellen Macht erhoben werden.

Poverty and Social Impact Analysis (PSIA) und Poverty Impact Assessment (PIA)

CHRISTIAN BERG / CLAUDIA GOTTMANN

PSIA ist ein systematischer Ansatz zur Analyse der Verteilungs- bzw. Wohlfahrtswirkungen von Politikreformen, der bis 2002 von der Weltbank entwickelt, seither mit zahlreichen methodischen Leitfäden und Trainingsmaterialien hinterlegt und inzwischen in einer ganzen Reihe von Ländern und Sektoren angewendet wurde (siehe <http://www.worldbank.org/psia>). Der PSIA-Ansatz fußt konzeptionell und methodisch auf Lernerfahrungen aus ökonomischen Analysen – z. B. Kosten-Nutzen-Analyse von Projekten und Investitionen, Ernährungspolitikanalyse – wie auch aus anthropologischen-soziologischen Analyseverfahren, speziell Social Impact Assessment (SIA).

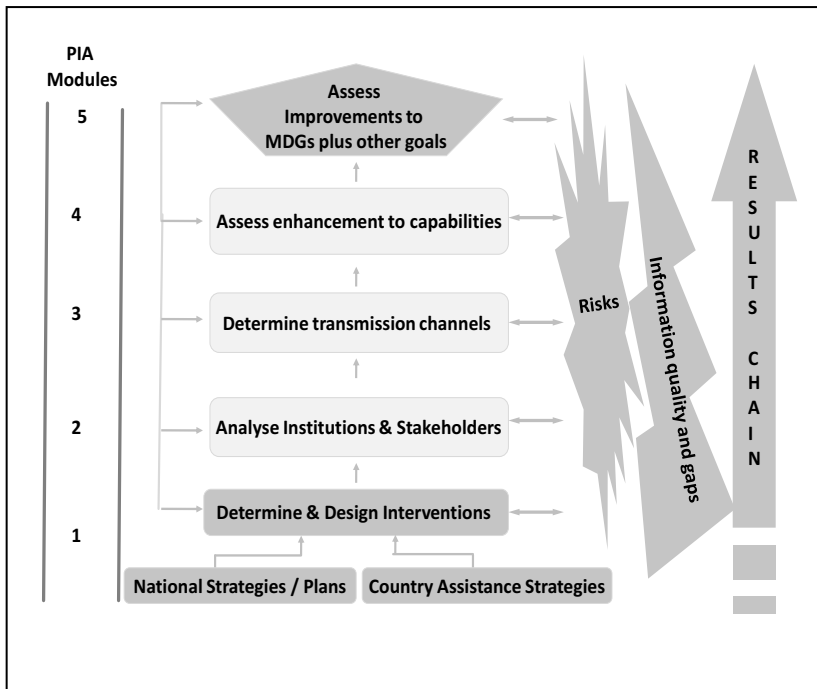
PIA wurde 2005/6 unter Beteiligung verschiedener Geber im Rahmen des OECD-DAC entwickelt und ist somit ebenfalls ein international anerkannter Analyseansatz zur Abschätzung der direkten und indirekten, positiven und negativen Wirkungen von Entwicklungsmaßnahmen. Anwendungserfahrungen gibt es für die Sektoren städtische Infrastruktur, Wasser/Abwasser, soziale Sicherung, ländliche Entwicklung, natürliches Ressourcenmanagement, Transport und Gesundheit (siehe http://www.oecd.org/document/62/0,3343,en_2649_34621_38609662_1_1_1_1,00.html).

Zwischen beiden Ansätzen (PSIA und PIA) besteht ein fließender Übergang und beide Ansätze basieren auf ähnlichen konzeptionellen Grund-

lagen. Jedoch wurde PIA eher mit Blick auf die mittlere Ebene (Programme und Subsektor-Strategien) entwickelt und PSIA eher mit Blick auf Politikreformen. Bei der Bewertung von komplexen Sektorprogrammen und -reformen kommt der PIA-Ansatz an seine Grenzen und es sollten zusätzlich Instrumente des PSIA-Ansatzes zum Einsatz kommen. Beide Ansätze weisen eine Reihe von Gemeinsamkeiten auf:

- PSIA wie PIA wurden für die Vorab- (ex-ante-) Analyse von Wirkungen, d. h. die Abschätzung zukünftiger Zustände, entwickelt. Ein Einsatz der Verfahren für die Wirkungserfassung nach Maßnahmen- oder Reformbeginn – im Verlauf, am Ende, ex-post – ist zwar in den jeweiligen Handbüchern vorgesehen, aber bislang nicht (PSIA) oder nur ansatzweise (PIA) erfolgt. So wurde PIA in der praktischen Anwendung bereits sowohl bei Projektfortschrittskontrollen erfolgreich angewandt und Elemente des Verfahrens sind im Kontext von Länderprogrammevaluierungen zum Einsatz gekommen.
- In beiden Ansätzen werden die Rollen und (möglichen) Veränderungsbeiträge und Widerstände von formellen und informellen Institutionen / Regelwerken und Organisationen analysiert.
- Beide Ansätze untersuchen die Verteilungs- bzw. Wohlfahrtswirkungen auf unterschiedliche gesellschaftliche Gruppen unter besonderer Berücksichtigung der armen und von Verarmung bedrohten Bevölkerungsgruppen.
- Die Analyse von Risiken für den Maßnahmen- bzw. Reformerfolg ist integraler Bestandteil von PIA wie auch – etwas weniger durchgängig – von PSIA.
- Ein entscheidender Analyseschritt in beiden Verfahren ist die Identifizierung von „Transmissionskanälen“, durch die Wirkungen erzielt werden sollen: (1) Beschäftigung; (2) Preise (Produktions- und Konsumpreise, Löhne); (3) Zugang(smöglichkeiten) zu öffentlichen und privaten Gütern und Dienstleistungen; (4) Steuern und Subventionen; (5) Verfügung über physisches, natürliches, humanes, finanzielles und soziales Kapital (*assets*); und – bei PIA – (6) Machtstrukturen (*authority*).

ABBILDUNG 21: Poverty Impact Assessment (PIA) Module



Ein wesentlicher Unterschied zwischen den Verfahren besteht im Aufwand für die Durchführung: Während eine PSIA ein Team von 10-25 Personen über mehrere Monate beschäftigt, um die relevanten Analyseschritte (einschließlich systematischer Datenerhebung) zu vollziehen, beruht PIA auf subjektiven Einschätzungen, die i. d. R. im Rahmen von Workshops eingeholt werden (sowie, falls vorhanden, auf der Verwendung von Sekundärdaten, Studien und Literatur), und wird von einem kleinen Team in einem Zeitrahmen von maximal 2 Wochen durchgeführt. Wichtige inhaltliche Unterschiede zwischen den Ansätzen sind:

- Im PIA-Modul 4 (siehe Abbildung) wird die Wirkung der Entwicklungsmaßnahme auf die Stärkung der wirtschaftlichen, menschlichen, politischen und soziokulturellen Fähigkeiten der Beteiligten (vor allem der Zielgruppen) sowie ihrer Sicherheit abgeschätzt (*capabilities*), differenziert nach den verschiedenen sozialen Gruppen. PSIA enthält nicht ausdrücklich einen solchen Analyseschritt.

- PIA baut auf der Nutzung vorhandener Daten auf, während PSIA eher eigene Datenerhebungen vorsieht und mit (makro-)ökonomischen Daten gespeiste Simulationsmodelle – allgemeine und partielle wirtschaftliche Gleichgewichte, Angebots-Nachfrage-Modelle, soziale Bilanzierungs-Matrizen, Verhaltenssimulationen auf Haushalts-ebene, usw. – nutzt, um die Wirkung von Politikreformen auf Armut abzuschätzen.

Eine besondere Stärke von PIA ist, dass es sich um ein gemeinsam von verschiedenen Gebern entwickelten Ansatz handelt und damit ein Schritt zur Harmonisierung von Instrumenten und Verfahren gemacht wurde. Da PIA nicht den Interessen bzw. methodischen Ansätzen nur eines einzelnen Gebers entspricht, ist das Verfahren auch für die Partner(regierungen) attraktiv.

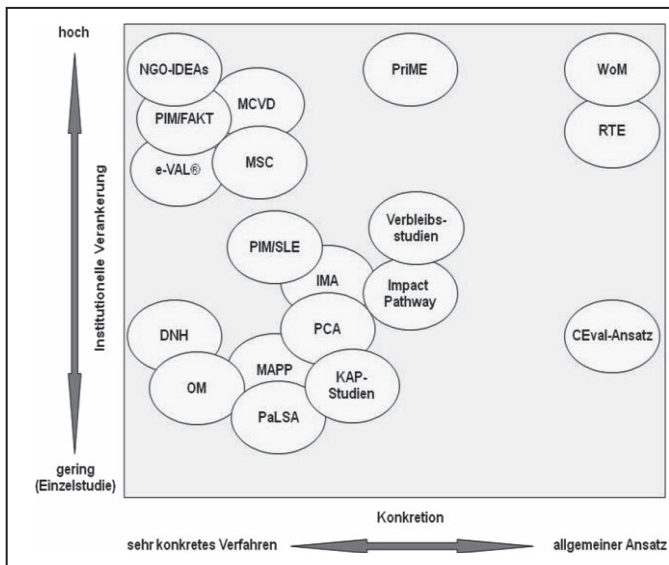
3. Synthese – Verfahren der Wirkungsanalyse in der Zusammenschau

In diesem Kapitel werden auf der Grundlage der Beschreibungen in Abschnitt 2.2 wesentliche Gemeinsamkeiten und Unterschiede zwischen den verschiedenen Wirkungsanalyseverfahren dargestellt (3.1) sowie bestehende Verbesserungsmöglichkeiten in den Verfahren bzw. ihren Beschreibungen identifiziert (3.2). Am Ende findet sich ein tabellarischer Überblick, in dem wichtige Merkmale der Verfahren zusammengestellt sind, die es erlauben, deren Eignung für einen konkreten Zweck einzuschätzen (3.3).

3.1 Gemeinsamkeiten und Unterschiede in den Verfahren

Grad der Konkretion und institutionellen Verankerung

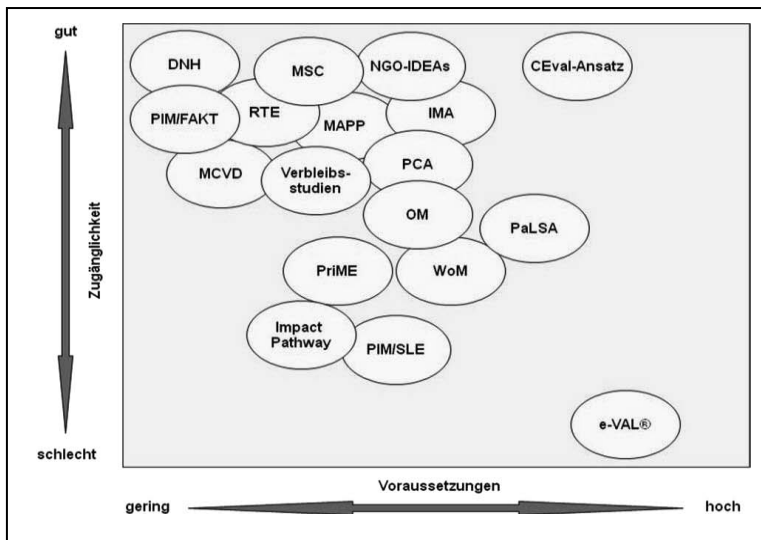
Ein großer Teil der Verfahren gibt konkrete Instrumente vor, wie Wirkungen erfasst werden sollen. Andere Verfahren stellen allgemeine Ansätze dar, die innerhalb eines konzeptionellen Rahmens in Bezug auf die konkreten Instrumente flexibel sind. Mischformen legen innerhalb eines allgemeinen Ansatzes einzelne Instrumente fest.



Einige Verfahren sind geeignet, in einer einzeln beauftragten Studie (z. B. Evaluierung) zur Wirkungserfassung eingesetzt zu werden. Diese sind auch geeignet für Förderorganisationen, die solche Studien zur Überprüfung des Erfolgs oder zum Lernen für andere Projekte oder Programme einsetzen wollen. Andere Verfahren sind von vornherein darauf angelegt, institutionell im Monitoring einer Organisation verankert und/oder kontinuierlich oder regelmäßig wiederkehrend durchgeführt zu werden. Sie generieren Wirkungsinformationen im Verlauf einer Entwicklungsmaßnahme. Viele Verfahren, die bisher überwiegend als Einzelstudien durchgeführt werden, sind allerdings prinzipiell auch für eine institutionelle Verankerung geeignet.

Zugänglichkeit und Voraussetzungen

Die Zugänglichkeit der Verfahren für Nutzer/-innen (Organisationen und Gutachter/-innen) ist sehr unterschiedlich. Manche sind gut und ausführlich beschrieben, mit Beispielen belegt und frei zugänglich (z. B. über das Internet und Publikationen). Andere Verfahren sind lizenziert oder organisationsintern dokumentiert und für einen externen Kreis so gut wie gar nicht verfügbar. Wieder andere sind nur unzureichend beschrieben oder publiziert, so dass prinzipieller Zugang zwar gegeben ist, dieser aber schwierig erreicht werden kann.



Die Voraussetzungen, die zur Anwendung eines Verfahrens benötigt werden, sind ebenfalls sehr unterschiedlich. Sie beschreiben in erster Linie die notwendige fachliche und methodische Expertise und den finanziellen Bedarf und/oder zeitlichen Aufwand für die Durchführung. Während die meisten Verfahren relativ voraussetzungsarm sind, sind andere eher voraussetzungsreich.

Anwendbarkeit für bestimmte Sektoren, Maßnahmenebenen und -typen

Die überwiegende Zahl der Verfahren ist in bestimmten entwicklungspolitischen (Teil-)Sektoren besonders erprobt worden, einige darunter eignen sich grundsätzlich auch für eine Anwendung in anderen Sektoren. Knapp die Hälfte der Verfahren wird demgegenüber insgesamt wenig sektorspezifisch angewendet oder ist – nach bisheriger Einschätzung – gänzlich sektorunabhängig einsetzbar (siehe Tabelle nächste Seite).

Eine ganze Reihe der vorgestellten Verfahren eignet sich für die Erfassung der Wirkungen von Entwicklungsmaßnahmen unabhängig davon, ob diese auf der lokalen (Mikro-)Ebene, der regionalen (Meso-)Ebene oder der nationalen (Makro-)Ebene ansetzen. Hierzu gehören der CEval-Ansatz, DNH, e-VAL[®], Impact Pathway, MSC, MCVD, PriME von InWent, Verbleibsstudien und WoM der GTZ. Andere Verfahren können auch auf institutioneller Ebene ansetzen und hierbei eine Grundlage für ein Change Management darstellen, z. B. KAP-Studien oder MAPP.

(Teil-)Sektoren, in denen das Verfahren besonders erprobt ist*	Verfahren (in Klammern: weitere Ansätze aus Abschnitt 2.3)
Bildung [110]	CEval-Ansatz; MSC; PriME; Verbleibsstudien
Gesundheitswesen [120]	CEval-Ansatz; KAP-Studien; MSC; PIM/FAKT; PriME
Wasserversorgung und Abwasser [140]	CEval-Ansatz; MAPP; PriME; (RSA)
Staat und Zivilgesellschaft allgemein [151]	IMA (für Gemeindeentwicklung); KAP-Studien (für Gemeindebeteiligung); MAPP; MCVD; PriME
Krisenprävention und Konfliktlösung [152]	MCVD; PCA; PriME
Sonstige soziale Infrastruktur und Dienste [160]	MSC; Verbleibsstudien (für Beschäftigungsförderung); PriME
Transport [210]	PriME
Radio, Fernsehen, Printmedien [22030]	OM
Informelle und halbformelle Finanzintermediäre [24040]	NGO-IDEAs, PIM/SLE
Privatwirtschaftliche und andere Dienste [250]	Impact Pathway (für Privatsektorentwicklung); PriME
Landwirtschaft [311]	IMA; KAP-Studien; MAPP; NGO-IDEAs; PIM/FAKT; PIM/SLE; PaLSA; PriME
Ressourcenmanagement/Umweltschutz [410]	CEval-Ansatz; IMA; Impact Pathway; MAPP; MCVD; PriME
Ländliche Entwicklung [43040]	CEval-Ansatz; IMA; MAPP; MSC; NGO-IDEAs; PIM/FAKT; PIM/SLE; PaLSA; PriME
Ernährungssicherung [520]	(HFES); IMA; KAP-Studien; NGO-IDEAs; PIM/FAKT; PIM/SLE; PriME; (RNA)
Humanitäre Hilfe [700]	CEval-Ansatz; (HFES); PIM/SLE; RTE; (RNA)
Insgesamt wenig sektorspezifisch oder ganz sektorunspezifisch angewendet	CEval-Ansatz; DNH; e-VAL [®] ; Impact Pathway; MAPP; MCVD; MSC; OM; (PIA); PriME; (PSIA); WoM

* Zahlen in Klammern weisen auf den entsprechenden Förderbereichsschlüssel nach DAC hin.

Demgegenüber stehen Verfahren, die sich überwiegend für die Wirkungsanalyse von Maßnahmen auf der Mikroebene – wie PIM/FAKT – oder, mindestens ansatzweise, zusätzlich auf der Mesoebene eignen: IMA, MAPP, OM, NGO-IDEAs, PIM/SLE, PaLSA, RTE.

PCA wurde für die Wirkungsanalyse von Maßnahmen auf der Meso- und Makroebene entwickelt.

Die Zahl geeigneter Wirkungsanalyseverfahren nimmt mit zunehmender Komplexität der zu untersuchenden Entwicklungsmaßnahme (Einzelprojekt → Einzelprogramm/Maßnahmenpaket → Sektorprogramm → Länder-/multisektorales Programm) ab bzw. je komplexer die Maßnahmen sind, desto eher eignet sich ein Mix von Methoden. Lediglich der CEval-Ansatz, DNH und e-VAL[®] – und ansatzweise MSC – nehmen für sich in Anspruch, für sämtliche Maßnahmentypen (einschließlich programmorientierter Gemeinschaftsfinanzierung bzw. Budgethilfe) anwendbar zu sein. Für die Wirkungsanalyse von Politiken und politikbeeinflussenden Maßnahmen werden derzeit methodische Ansätze entwickelt und erprobt, hier gibt es noch keine Standards (siehe Abschnitt 2.3).

Eignung für die Erfassung verschiedener Arten von Wirkung

Sämtliche in den Steckbriefen dokumentierten Verfahren erfassen die direkten (geplanten) Wirkungen einer Entwicklungsmaßnahme infolge der Nutzung ihrer Leistungen/Outputs durch Mittler und/oder Zielgruppen, wie sie für eine einfache Überprüfung der Zielerreichung von Vorhaben erforderlich ist. Hinsichtlich der Erfassung sonstiger Wirkungen lässt sich Folgendes feststellen:

Indirekte (positive und negative) Wirkungen (<i>impact</i> nach OECD/DAC) können erfassen:	Unerwartete (positive und negative, direkte oder indirekte) Wirkungen können erfassen:
Uneingeschränkt: CEval-Ansatz, DNH, IMA, MAPP, MSC, NGO-IDEAs, PaLSA, PIM/FAKT und /SLE, MCVD, PCA, WoM	Uneingeschränkt: CEval-Ansatz, DNH, IMA, MAPP, MSC, PaLSA, PCA (v. a. bei Verwendung von DNH), MCVD, WoM
Eingeschränkt: e-VAL [®] (Erfassung nur sofern Wirkungen thematisiert werden), Impact Pathway	Eingeschränkt: e-VAL [®] (Erfassung nur sofern Wirkungen thematisiert werden), Impact Pathway, KAP-Studien (da nur direkte Wirkungen), NGO-IDEAs, PIM/FAKT und /SLE (da keine systematische Einbeziehung externer Perspektive oder von Kontrollgruppen), RTE, WoM

Besonders spezialisiert auf die Analyse ungeplanter Wirkungen sind DNH, MAPP, MSC, PaLSA und PCA (bei gleichzeitiger Verwendung von DNH).

Alle vorgestellten Wirkungsanalyseverfahren sind naturgemäß grundsätzlich geeignet, Wirkungen von Entwicklungsmaßnahmen auf der Ebene der Bevölkerung zu erfassen, d. h. von Maßnahmen, die mittelbar oder unmittelbar die Stärkung der Handlungskompetenz von Individuen, Haushalten oder Gruppen/Gemeinden zum Ziel haben. Geht es darüber hinaus um die Erfassung von Wirkungen auf der Ebene der Organisation, d. h. von Maßnahmen zur Stärkung der Leistungsfähigkeit von öffentlichen und privaten Dienstleistern, Verbänden, Netzwerken usw., bieten sich v. a. (aber nicht ausschließlich) der CEval-Ansatz, DNH, Impact Pathway, OM, MAPP, MCVD, PCA, PriME von InWEnt und WoM der GTZ an. Zur Analyse von Wirkungen auf der Systemebene, d. h. von Maßnahmen zur Verbesserung der Handlungen und Entscheidungen auf politischer Ebene, können gemäß der Darstellung in den entsprechenden Steckbriefen v. a. der CEval-Ansatz, Impact Pathway, PCA und WoM der GTZ dienen (daneben auch neuere methodische Ansätze zur Messung von Politikbeeinflussung, siehe Abschnitt 2.3).

Verfahren	Erfassung von Wirkungen auf*:		
	Bevölkerungsebene	Organisationsebene	Systemebene
CEval-Ansatz	—————	—————	—————
DNH	—————	—————	
e-VAL [®]	-----	-----	-----
IMA	—————		
Impact Pathway	—————	—————	—————
KAP-Studien	—————	-----	
MAPP	—————	—————	-----
MSC	—————	—————	-----
NGO-IDEAs	—————		
OM	—————	—————	
PIM/FAKT	—————	-----	
PIM/SLE	—————	-----	-----
PaLSA	—————		
MCVD	—————	—————	-----
PCA	—————	—————	—————
PriME	-----	—————	
RTE	—————		
Verbleibsstudien	—————		
WoM	—————	—————	—————

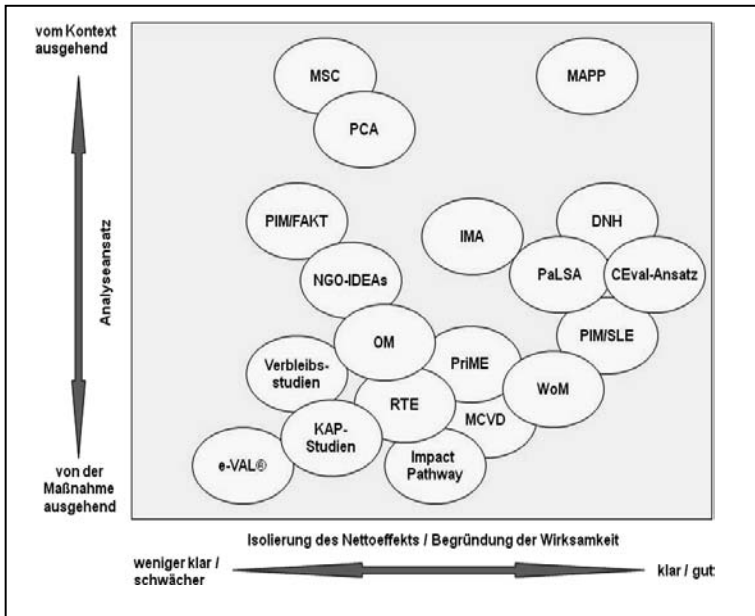
* Gestrichelte Linien bedeuten eine eingeschränkte (d. h. nur unter Modifikation des Verfahrens und/oder nur unter bestimmten Bedingungen oder Fragestellungen mögliche) Eignung.

Ausgangspunkt der Analyse und Begründung der Wirksamkeit

Während eine Gruppe von Verfahren analytisch von der Entwicklungsmaßnahme ausgeht, oft anhand der Formulierung von Wirkungsketten, ausgehend vom Input, über die Outputs, Outcomes bis hin zum anvisierten Impact auf den unterschiedlichen Ebenen, geht eine andere Gruppe von Verfahren eher vom Kontext aus, d. h. von den Veränderungen, die sich in einem Zeitraum ergeben haben (auf der Bevölkerungsebene, Organisations- oder

Systemebene) und ordnet erst dann diese Veränderungen verschiedenen Ursachen (Entwicklungsmaßnahme und externen Einflüssen) zu.

Während festgestellte Veränderungen „Bruttowirkungen“ darstellen, da sie sämtliche Einflüsse mit beinhalten, ist die Wirkung der Entwicklungsmaßnahme der „Nettoeffekt“, da er eindeutig auf die zu analysierende Maßnahme zurückgeführt werden kann. Dies bedeutet, sie ist von anderen Einflüssen bereinigt. Die Verfahren unterscheiden sich darin, ob diese Zuschreibung des Nettoeffekts auf die Maßnahmen explizit geschieht, mit welcher Validität und wie gut sie quantitativ und/ oder qualitativ begründet wird.



Verwendung quantitativer und qualitativer Methoden

Die Frage, ob Verfahren qualitativen oder quantitativen Charakter haben, muss jeweils an verschiedenen Stellen der Analyse beantwortet werden, und zumeist handelt es sich um Mischformen. So kann ein Verfahren grundsätzlich qualitativer Natur sein, indem z. B. dem Zuordnungsproblem mit Hilfe von Einschätzungen der Zielgruppen begegnet wird, und dennoch quantitative Elemente enthalten (wie z. B. Punktwertungen) sowie mit großen Stichproben arbeiten, so dass sich die Ergebnisse der Einzelerhebungen aggregieren lassen (wie z. B. bei IMA, MAPP, MSC, NGO-IDEAs, PaLSA). Hier handelt es sich somit um qualitativ-quantitative Mischformen,

welche in den letzten Jahren immer stärkeren Eingang in die Praxis gefunden haben. Dies kann als ein Zeichen dafür gesehen werden, dass der Paradigmenstreit um Methoden im Großen und Ganzen vorüber ist. Stattdessen wird pragmatisch sowie inhaltlich mit der Frage umgegangen, ob eine quantitative oder qualitative Herangehensweise anzuwenden ist.

Art der Validierung von Wirkungsdaten/-informationen

Die Verfahren verfolgen unterschiedliche Ansätze in Bezug auf die Frage, welche Art der Validierung von Daten sie verwenden. Dabei sind die Grenzen zwischen den Ansätzen fließend.

- a) Kombination diverser Instrumente und Perspektiven:
Einige Verfahren verwenden – systematisch angelegt – verschiedene Erhebungsinstrumente und greifen auf die Perspektiven mehrerer Akteure zurück, allerdings meist ohne festzulegen, welche Instrumente und Methoden konkret anzuwenden seien. Zu dieser Gruppe gehören der CEval-Ansatz, IMA, OM, NGO-IDEAs (in PANgo), PriME von InWEnt (das in bestimmten Phasen eines Projektes Fragebögen mit festgelegten Fragen vorschreibt), RTE und WoM der GTZ.
- b) Zeitgleiche Berücksichtigung verschiedener Akteursperspektiven:
Bei einigen Verfahren besteht die Validierung zentral darin, dass sie zeitgleich mehrere Akteursperspektiven einbeziehen (in der Regel in Form eines Workshops) und/oder auf Diskursen beruhen: Mehrere Akteure kommen zusammen, stellen ihre Sichtweisen dar und diskutieren sie, oft unter Verwendung festgelegter (und ggf. unterschiedlicher) Instrumente, die ein zweites Mittel der Validierung sind, indem sie Gegenprüfungen ermöglichen. Hierzu zählen CEval-Ansatz, DNH, MAPP und PaLSA.
- c) Zeitlich versetzte Berücksichtigung mehrerer Akteursperspektiven:
Andere Verfahren nutzen eine zusätzliche Form der Triangulation, indem sie zwischen unterschiedlichen Akteursgruppen (z. B. Basisgruppe, NRO, Geldgeber) differenzieren. Jede Akteursgruppe geht für sich vor wie in Typ b). Die Ergebnisse werden dann zeitversetzt miteinander abgeglichen. Hierzu zählen Impact Pathway, PIM/FAKT und PIM/SLE (die auch Treffen zwischen den Akteursgruppen vorsehen) und MCVD.
- d) Aggregation individueller Akteursperspektiven:
e-VAL[®], MSC und Verbleibsstudien greifen zunächst auf die Darstellung/Messung von Individuen zurück. Bei MSC und Verbleibs-

studien bleibt die Option, deren Darstellung mit anderen Instrumenten zu validieren, dies ist aber nicht systematisch vorgesehen. e-VAL[®] sieht eine computerisierte Interpretation der Interviews vor, die von den Befragten erneut überprüft wird. Die Erstellung einer Statistik / ökonomischen Untersuchung bei quantitativen Verfahren oder eine Triangulation bei qualitativen Verfahren erfolgt dann bei der Aggregation: Viele individuelle Sichtweisen, deren Tatsachengehalt nicht überprüft ist, werden miteinander abgeglichen, um Aussagen über ein gesamtes Programm oder Teile davon zu machen. Eine Triangulation mit weiteren Methoden ist nicht Inhalt dieser Verfahren, wenn sie auch oft ausdrücklich gewünscht wird.

Möglichkeit der Aggregation von Wirkungsdaten/-informationen

Die Aggregation von Daten bzw. Informationen ermöglicht die Zusammenfassung der Ergebnisse von Untersuchungen auf lokaler Ebene (Individuen, Haushalte, Gruppen, Gemeinden) zwecks Darstellung der Situation oder Veränderung auf regionaler Ebene (Kreis, Departement, Provinz usw.) oder nationaler Ebene. Dies ist möglich bei der Anwendung quantitativer Methoden oder – wenn auch in abgeschwächter Form – von Punktbewertungen. Eine notwendige Voraussetzung ist die Verwendung gleicher Indikatoren. Werden im Rahmen einer stark partizipativen Vorgehensweise die Indikatoren lokal von der Zielgruppe mitbestimmt, so müssen zumindest allgemein verbindliche Schlüsselindikatoren in allen Untersuchungen verwendet werden, um eine Aggregation der Ergebnisse zu ermöglichen.

Etwa die Hälfte der Verfahren erlaubt eine Aggregation der Daten, während bei den übrigen eine Aggregation nicht oder nur eingeschränkt möglich oder nicht vorgesehen ist bzw. von den gewählten Indikatoren abhängt (siehe Überblick in Abschnitt 3.3).

Eignung für die Erfassung von Zuständen und Prozessen

Die dargestellten Verfahren unterscheiden sich darin, inwieweit sie einen Schwerpunkt auf die Erfassung der Veränderung (oder auch Stagnation) von Zuständen oder auch auf die Erfassung von Prozessen legen, d. h. zusätzlich die Frage beantworten, wie und warum Wirkungen erzielt werden. Geht es bei der Zustandsanalyse um den Vergleich von Situationen zu verschiedenen Zeitpunkten (entweder über zeitlich versetzte Messungen oder nachträgliche Rekonstruktion des Ausgangszustands vor Beginn einer Entwicklungsmaßnahme), werden bei der Prozessanalyse der Ablauf von Ent-

wicklungen sowie Interaktionen (zwischen Akteuren und Institutionen) untersucht.

Einen Schwerpunkt auf die Erfassung von Zuständen legen:	Zustände und Prozesse gleichermaßen erfassen:	Einen Schwerpunkt auf die Erfassung von Prozessen legen:
e-VAL [®] (sofern nicht auch Prozesse thematisiert werden), KAP-Studien, NGO-IDEAs	CEval-Ansatz, IMA, Impact Pathway, PIM/FAKT und /SLE, PaLSA, PCA, PriME von InWEnt, Verbleibsstudien, WoM der GTZ	DNH, MAPP, MSC, OM (eingeschränkt), MCVD, RTE

Grad der Beteiligung von Trägern und Zielgruppen an der Wirkungsanalyse

Die Rolle, die den Entwicklungsmaßnahmen durchführenden Organisationen (Trägern) und den Zielgruppen (Endbegünstigten) in der Planungs-, Durchführungs-, Analyse- und Bewertungsphase einer Wirkungsuntersuchung zugewiesen wird, reicht in den hier vorgestellten Verfahren von passiver Beteiligung (als Informationsgeber) bis zur gänzlich eigenverantwortlichen Durchführung (im Rahmen des projektinternen Monitoring).

Einerseits ist die Unabhängigkeit (und damit größere Objektivität) von Wirkungsstudien – etwa im Rahmen externer Evaluierungen – ein wichtiger Gütestandard. Andererseits ist z. B. zu beachten, dass Verfahren, die auch für das kontinuierliche Monitoring gestaltet sind, nicht nur Wirkungen erfassen, sondern auch Einfluss auf die Wirkung ausüben wollen. Sie sollen zur kontinuierlichen Verbesserung des Programms beitragen und selbst direkte Auswirkungen haben, auch auf die Managementkapazität von Trägern und Zielgruppenorganisationen (IMA, MSC, NGO-IDEAs, OM, PIM, PriME, RTE etc.). Eine Absicht dahinter ist: Wenn ein Verfahren für den Projekterfolg und die Zielgruppen hilfreich ist, dann hat es eine größere Chance, auch verlässlich durchgeführt und dokumentiert zu werden. Das Problem vieler Wirkungsstudien, dass nicht genügend Daten über Wirkungen im Projektverlauf gesammelt wurden, verringert sich dadurch. Indem diese Verfahren sich darauf konzentrieren, kontinuierlich die Sicht der Zielgruppen und anderer Beteiligter zu erfassen, welche Wirkungen eine Entwicklungsmaßnahme zeitigt, tragen sie gerade dazu bei, dass in diesen Vorhaben mehr und relevantere Wirkungsdaten erfasst werden als in anderen.

Verknüpfungen und Weiterentwicklungen von Verfahren

Manche der vorgestellten Wirkungsanalyseverfahren bauen auf anderen Verfahren auf, die bereits früher entwickelt wurden:

- OM und MSC sind – ohne einschneidende Adaptionen – in MCVD enthalten, MAPP in PaLSA, DNH (in der Regel) in PCA.
- PIM/SLE ist eine Weiterentwicklung von PIM/FAKT in Richtung einer stärkeren Repräsentativität der Erhebung und Berücksichtigung des Kontrafaktums (durch Vergleichsgruppen), PIA eine Anpassung von PSIA in Richtung eines einfacheren und kostengünstigeren Verfahrens.

Dies, wie auch die Tatsache, dass viele der Verfahren auf unterschiedlichste Ansätze der empirischen Sozialforschung, auf Methoden des Participatory Rural Appraisal (PRA) oder (wie IMA und PaLSA) auf Ansätze der Systemanalyse zurückgreifen bzw. diese Ansätze und Methoden modifizieren und sie ggf. in eine andere Ordnung stellen, zeigt, dass Wirkungsanalyseverfahren keineswegs statisch sind, sondern jederzeit auf bestimmte Anforderungen (etwa Kostengünstigkeit), Gütestandards oder Organisationskulturen hin angepasst und verändert werden können.

3.2 Anwendungs- und Verbesserungsmöglichkeiten

In diesem Handbuch wurde bewusst darauf verzichtet, einzelne Verfahren als generell besser oder schlechter zu bewerten. Diese Entscheidung gründet auf der Einsicht, dass jedes der hier vorgestellten Verfahren zum Einsatz geeignet sein kann, je nachdem welche weiteren Randbedingungen bestehen oder welche Vorgaben erfüllt werden sollen. Die richtige Entscheidung hängt dabei in erster Linie vom Maßnahmentyp und Zielkonzept der Entwicklungsmaßnahme ab, wie im Eingangskapitel und in den Steckbriefen auch an mehreren Stellen betont wird. Daneben sind die mit der Datenerhebung verbundenen übergeordneten Ziele und die vorhandenen Kapazitäten (Fachlichkeit, Zeit und Geld) für die Eignung eines Verfahrens wesentlich. Die unter dem Begriff Wirkungsanalyse zusammengefassten Untersuchungsmaßnahmen reichen von 1-2 wöchigen Evaluierungen bis hin zu mehrmonatigen Wirkungsstudien, an die sicher nicht die gleichen Ansprüche gestellt werden können. Systeme des Wirkungsmonitoring müssen wieder andere Ansprüche erfüllen.

Aus rein inhaltlicher Sicht ist die Analyse bis hin zu den Systemgrenzen der Entwicklungsmaßnahmen wünschenswert, unter Einbezug aller Stakeholder, auch der nicht von der Maßnahme unmittelbar betroffenen,

denn nur so kann das Mikro-Makro-Paradoxon (vgl. Abschnitt 1.1) letztlich aufgelöst werden.

Ideal wäre es, sowohl einen fundierten Wirkungsnachweis zu geben, mit dem rigorose Wissenschaftsstandards voll erfüllt werden, wozu die Anwendung quantitativer Methoden notwendig ist, als auch mit Hilfe qualitativer Methoden eine umfassende Erklärung für die Zusammenhänge und Prozesse zu erhalten, damit die Ergebnisse mit einem Lernerfolg verbunden werden können. Um quantitative Methoden sinnvoll einsetzen zu können, sind in der Regel qualitative Vorstudien erforderlich, da der Kontext gut bekannt sein muss. Auch eine möglichst weitgehende Erfassung relevanter, auch unerwarteter Veränderungen erfordert die Anwendung qualitativer Methoden. Deshalb ist es wichtig, die Prinzipien der EZ (inkl. Paris-Agenda) auch in Wirkungsanalysen zu berücksichtigen. Diese fordern zum Beispiel das *empowerment* durch Einbeziehung der Zielgruppen als aktive Akteure in eine Entwicklungsplanung und implizieren ein partizipatives Vorgehen.

Diese vielfältigen Ansprüche zusammen mit den immer komplexer werdenden Zielkonzepten und Maßnahmenbündeln der EZ sind nicht alle zu jedem Zeitpunkt und bei jeder Wirkungsanalyse erfüllbar. Sie erfordern daher eine klare methodische Entscheidung mit dem Mut zur Lücke (nach dem Motto: alles geht nicht), aber auch den kreativen Einsatz eines wohl durchdachten Methodenmix, der mit Hilfe eines angemessenen Budgets versucht, den jeweils wichtigsten Ansprüchen gerecht zu werden.

Die hier vorgestellten Verfahren stellen insgesamt ein reichhaltiges Reservoir dar, aus dem geschöpft werden kann. Auch wenn alle Verfahren praxisbewährt sind und daher empfohlen werden können, besteht im Einzelnen noch weiterer Verbesserungsbedarf. Hier werden drei Punkte herausgestellt:

Bessere Einbettung

Bei Anwendung verschiedener Verfahren sollten die empirisch erhobenen Daten mit bereits vorhandenen Daten aus früheren Untersuchungen (einschließlich Projektdaten) oder aus der lokalen oder nationalen Statistik systematischer – und erst nach kritischer Prüfung ihrer Qualität – verknüpft werden. Somit würden ein besserer *cross-check* und eine bessere Kontextualisierung erzielt und damit könnten Aussagen besser fundiert bzw. strukturelle Aussagen überhaupt gemacht werden. Darüber hinaus ließen sich dadurch Doppelerhebungen vermeiden.

Bessere Darstellung der Auswertungsmethodik

Weiterhin ist zu bemerken, dass es neben den zumeist gut dargestellten Erhebungsinstrumenten in vielen Verfahren an einer genaueren Beschreibung oder an einem Instrumentarium zur Auswertung und Darstellung der Ergebnisse fehlt. Auch der Umgang mit Widersprüchen oder Kontroversen aus unterschiedlichen Datenquellen wird noch zu wenig thematisiert. Im Sinne einer stärkeren Standardisierung wären hier ebenfalls Verbesserungen wünschenswert.

Bessere Erfassung der Kosten und des Zeitbedarfs

Zuletzt wäre es hilfreich, wenn die Dokumentation zu Kosten und Zeitbedarf der Verfahren verbessert würde, so dass deren Kalkulation im Vorfeld besser möglich wäre.

3.3 Überblick über Anwendungsmöglichkeiten und Voraussetzungen der Verfahren

In der folgenden Tabelle sind die verschiedenen Wirkungsanalyseverfahren so zusammengestellt, dass ihre Eignung für einen konkreten Zweck anhand wichtiger Merkmale schnell eingeschätzt werden kann. Für die Einschätzungen und Bewertungen der verschiedenen Merkmale zeichnen die jeweiligen Autor/-innen der Verfahrensbeschreibungen (siehe Kap. 2.2) verantwortlich.

Kriterium:	Anwendungsbereich ⁵³	Gesellschaftliche Ebene der Wirkung ⁵⁴	Erfasste Wirkungen		Eignung für:		Aggregation der Daten möglich	Besondere Qualifikationen ⁵⁵	Zugänglichkeit ⁵⁶
			Auch indirekte ⁵⁷	Auch unmittelbare wartete ⁵⁸	Interne M&E ⁵⁸	Einzelne Studien ⁵⁹			
Vorfahren:									
CEval- Ansatz zur Wirkungs- evaluation / Stockmann- scher Ansatz	Alle Sektoren, bisher in folgenden Bereichen erprobt: Entwicklungszusammenarbeit, Umwelt, Bildung, Berufsbildung, Gesundheit, ländliche Entwicklung, Wasser- und Abwasser, humanitäre Hilfe etc.; alle Interventionsarten, auf allen Ebenen (Mikro/Meso/Makro)	Bevölkerungs-, Organisations- und Systemebene	Ja	Ja	Ja	Ja	Ja	Besondere Methodenkompetenz für die Entwicklung des Analyserasters, zur Datenerhebung, -auswertung und -interpretation; <i>Reportingskills</i> , ggf. Moderationserfahrung	Gut (Ansatz u. a. erläutert in STOCKMANN (2006); Evaluation und Qualitätsentwicklung; englische, spanische, chinesische Übersetzung vorhanden, thailändische ist in Vorbereitung)
Do No Harm (DNH)	Alle Maßnahmen, die im Kontext gewaltsamer Konflikte durchgeführt werden (und deshalb potenziell Wirkungen auf den Konflikt haben), auf allen Ebenen (Mikro/Meso/Makro)	Vor allem Bevölkerungsebene und Organisationsebene	Ja	Ja	Ja	Ja	Nein	Keine	Gut
e-VAL [®]	Kann eingesetzt werden ohne sektorale Begrenzungen auf allen Maßnahmenebenen (Mikro/Meso/Makro) für alle Typen	Kann für Wirkungsanalysen genutzt werden	Ja	Ja	Ja	Ja	Eingeschränkt Ja, insbesondere bei Eingriffen	Die Interviews werden durch speziell geschulte Mitarbeiter durchgeführt	Eingeschränkt (Literatur, zertifizierte Gutachter, entsprechende Vorgehen werden vorausgesetzt)

⁵³ Nur Bereiche und Maßnahmenebenen (Mikro/Meso/Makro), in denen das Verfahren bislang mehrfach angewendet bzw. ausreichend erprobt wurde; prinzipielle Eignungen für andere Anwendungsbereiche in eckigen Klammern.

⁵⁴ Wirkungen auf **Bevölkerungsebene** (d. h. von Maßnahmen zur Stärkung der Handlungskompetenz von Individuen, Haushalten, Gruppen/Gemeinden), auf **Organisations-ebene** (d. h. von Maßnahmen zur Stärkung der Leistungsfähigkeit von Organisationen, z. B. öffentliche und private Dienstleister, Verbände, Netzwerke), auf **Systemebene** (d. h. von Maßnahmen zur Verbesserung von Handlungen und Entscheidungen auf politischer Ebene), auf der **Ebene der übergeordneten Wirkung** (d. h. von Maßnahmen zur zukunftsorientierten Gestaltung von Veränderungen in einer globalisierten Welt).

⁵⁵ Nur spezielle Qualifikationen. Vertrautheit mit dem Verfahren und allgemeine Fähigkeiten zur Moderation und zu partizipativem Vorgehen werden vorausgesetzt.

⁵⁶ Kriterien: Literaturtage, Qualifizierungsangebot, Notwendigkeit einer Zertifizierung, Bedarf für spezielle Software.

⁵⁷ Impact nach DAC/OECD (siehe Glossar), die Erfassung direkter geplanter Wirkungen wird vorausgesetzt.

⁵⁸ Im Rahmen des Project Cycle Management (PCM), auch geeignet für Selbstevaluierung (durch die durchführende Organisation / den Partner vor Ort).

⁵⁹ Z. B. im Rahmen einer externen Evaluierung, genutzt

Kriterium: Verfahren:	Anwendungsbereich ⁵³	Gesellschaftliche Ebene der Wirkung ⁵⁴	Erfasste Wirkungen		Eignung für:		Aggregation der Daten möglich	Besondere Qualifikationen ⁵⁵	Zugänglichkeit ⁵⁶
			Auch indirekte ⁵⁷	Auch unerwartete	Inter- netes M&E ⁵⁸	Einzelne Studien ⁵⁹			
	pen von EZ-Projekten und -Programmen; eignet sich zum Vergleich der Einschätzungen verschiedener Interessengruppen. Ebenso sind übergreifende Auswertungen auf Ebene von Sektoren, Multi-Gebieten, Programmen oder Länderportfolios möglich	den, wenn Wirkungen von den Interviewten thematisiert werden; dies bildet die Basis für weitergehende Wirkungsanalysen					schätzungen zu Themen, die schwer „objektiv“ messbar sind, sich aber subjektiv oft treffsicher beurteilen lassen, z. B. „ownership“ und „capacity development“	terviewer durchgeführt und durch dafür ausgebildete Interpretatoren ausgewertet	chende Software und Zugang zum GTZ-Server notwendig (wenig)
Impact Monitoring and Assessment (IMA)	Ländliche Entwicklung auf Mikro- und Mesoebene [analog anwendbar bei natürlichen Ressourcen und Gemeinwesenentwicklung]	Bevölkerungsebene und Naturraum	Ja	Ja	Ja	Ja	eher Ja (hängt von den ausgewählten Tools ab)	Verständnis für Kybernetik / Netzwerkstrukturen	Gut
Impact Pathway	Nicht sektorspezifisch und auf allen Ebenen (Mikro/Meso/Makro), angewendet z. B. für Privatssektorentwicklung, Forschung im Bereich natürlicher Ressourcen, Länderprogrammmonitoring	Bevölkerungs-, Organisations- und Systemebene	eher Ja	eher Ja	Ja	Ja	eher Nein	Keine	Gut
KAP-Studien	Maßnahmen, bei denen zielgruppenspezifische Veränderungen von Wissen, Einstellungen und Verhalten angestrebt werden, z. B. Gesundheit, Landwirtschaft, Gemeindebeteiligung, Frauenförderung etc.	Bevölkerungsebene	eher Ja	Ja	Ja	Ja	Ja (sofern im Design angelegt)	Erstellung und Auswertung von Fragebögen	Gut (Handreichung im Web, Standardfragebögen z. B. bzgl. sexueller reproduktiver Gesundheit verfügbar)

Kriterium:	Anwendungsbereich ⁵³	Gesellschaftliche Ebene der Wirkung ⁵⁴	Erfasste Wirkungen		Eignung für:		Aggregation der Daten möglich	Besondere Qualifikationen ⁵⁵	Zugänglichkeit ⁵⁶
			Auch indirekte ⁵⁷	Auch unerwartete	Inter- nes M&E ⁵⁸	Einzelne Studien ⁵⁹ die			
Verfahren:									
Method for Impact Assessment of Programmes and Projects (MAPP)	Alle Sektoren (mit Wirkungen auf der Bevölkerungsebene), geeignet zur Evaluierung komplexer Konzepte (Armutsminimierung, Demokratisierung etc.) und verschiedener Maßnahmenpakete v.a. auf Mikro- und Mesoebene	Bevölkerungsebene (kann prinzipiell auch Wirkungen auf Organisationsebene erfassen)	Ja	Ja	Nein (aber für Selbst-Evaluierung)	Ja	Ja, Aggregation der Punktewerte ebenso möglich wie Zusammenfassung der qualitativen Daten	Keine	Relativ gut, MAPP wurde mehrfach in Publikationen dokumentiert (s. auch Homepage von Susanne Neubert auf www.die-gdi.de)
Most Significant Change (MSC)	Ländliche Entwicklung, Bildung, Gesundheit; grundsätzlich geeignet für komplexe Maßnahmen sowie für Maßnahmen auf Mikro-, Meso- und Makroebene, die auf Armutsbekämpfung, sozialen Wandel, Bereitstellung sozialer Dienstleistungen abzielen	Bevölkerungsebene (prinzipiell auch für Organisations-, System- und übergeordnete Wirkungsebenen geeignet)	Ja	Ja	Ja	Ja	eher Ja (Verwendung thematischer Codierung bei der Datenanalyse mit anschließender Quantifizierung)	Keine	Gut (Handbuch im Internet; keine Zertifizierung oder spezielle Software erforderlich; Anwendergruppe mit E-Mail-Verteiler und Datenbank konstituiert seit 2000)
NGO-IDEAS Impact Toolbox	Selbsthilfe(förderung i. w. S. auf der Mikroebene [speziell für Mikrofinanz entwickelt, aber leicht übertragbar])	Bevölkerungsebene	Ja	eher Ja	Ja	Nein (zu testen)	eher Ja (falls gleiche Indikatoren)	Keine	Gut (pdf-Download bei www.ngo-ideas.net)
Outcome Mapping (OM)	Programme, Projekte und Organisationen in allen Sektoren - auch Medien - im ländlichen und urbanen Kontext (Mikro- und Mesoebene), M&E von Programmen zur Armutsbekämpfung	Bevölkerungs- und Organisationsebene	eher Nein (Fokus auf <i>outcome</i>)	eher Nein	Ja	Ja	Quantitativ und qualitativ (kontextbezogene 'Geschichten')	Keine (Training wird angeboten und Erfahrungsaustausch über website möglich)	Gut (Handbuch, detaillierte Materialien und Handrechen-Checklisten)
Participatory Impact Monitoring (PIM) / FAKT	Landwirtschaft / Ländliche Entwicklung i. w. S. (prinzipiell auch in anderen Sektoren anwendbar, z.B. Ernährung, Gesundheit); zielgruppenmaße Projekte (Mikroebene) aller Art, die auf Selbsthilfe abheben]	Bevölkerungsebene (prinzipiell auch für Organisationsebene geeignet)	Ja	eher Ja (Befürchtungen der Beteiligten)	Ja	Nein	eher Nein	Keine	Gut (pdf-Download bei www.fakt-consult.de in E, F, S, P und Brasilianisch)

Kriterium:	Anwendungsbereich ⁵³	Gesellschaftliche Ebene der Wirkung ⁵⁴	Erfasste Wirkungen		Eignung für:		Aggregation der Daten möglich	Besondere Qualifikationen ⁵⁵	Zugänglichkeit ⁵⁶
			Auch indirekte ⁵⁷	Auch unerwartete	Inter- netes M&E ⁵⁸	Einzelne Studien ⁵⁹ die			
Verfahren:									
Participatory Impact Monitoring (PIM) / SLE	Landwirtschaft / Ländliche Entwicklung i. w. S. (prinzipiell auch in anderen Sektoren anwendbar, z.B. Ernährung, Gesundheit); zielgruppenmaße <i>Livelihood</i> -Projekte auf der Mikro- (eingeschränkt auch Meso-) Ebene aller Art (Ressourcenschutz, Einkommen schaffende Maßnahmen, Spar- und Kreditgruppenförderung etc.)	<u>Bevölkerungs-</u> ebene (prinzipiell auch für Organisations- und Systemebene geeignet)	Ja	eher Ja (Be-fürch-tungen der Be-teiligten)	Ja	Ja	eher Nein (Quantifizierung zwar an-geleitet, Indika-toren sind aber oft lokal spezi-fisch)	Keine	Eingeschränkt (SLE-Publikation vergriffen, Hand-buch allerdings bei Weltungerhilfe verfügbar)
Participatory Livelihood System Analysis (PaLSA)	Maßnahmen auf der Mikro- und Mesoebene in den Sektoren Ländliche Entwick-lung (inkl. einkommenschaffende Maß-nahmen) und Landwirtschaft; speziell geeignet für systemische Untersuchungen	<u>Bevölkerungs-</u> ebene	Ja	Ja	Ja	Ja	Ja	Verständnis für Kybernetik; bei Übersetzung hö-herer fachlicher Anspruch an Ü-bersetzer	Gut, allerdings bis-her wenig geschul-tes Personal vor-handen
Participatory Methodology for Measuring the Contribution of Volunteering to Development (MCVD)	Armutsreduzierung, Umweltschutz, Re-gierungsführung, Jugend, Krisenbewälti-gung und -prävention (grundsätzlich alle Sektoren, in denen Freiwillige einen Ent-wicklungsbeitrag leisten) auf allen Maß-nahmenebenen (Mikro/Meso/Makro)	<u>Bevölkerungs-</u> und <u>Organis-</u> ationsebene, <u>Systemebene</u> (prinzipiell können auch Wirkungen auf Systemebene erfasst werden)	Ja	Ja	Ja (be-sonders auch für Selbst-evaluierung ge-eignet)	Ja	Ja (aufwändig, da qualitative Daten; zudem abhängig von der Qualität der Dokumen-tation von Workshops auf Projektebene)	Keine	Gut (Handbuch zur Methode wird der-zeit überarbeitet, anschließend auf http://www.unv.org frei verfügbar)
Peace and Conflict Assessment (PCA)	Alle Interventionen, die auf Konflikte einwirken sollen (<i>working on conflict</i>) oder das Potential dafür haben (<i>working in conflict</i>), speziell solche auf Meso- und Makroebene	<u>Bevölkerungs-</u> , <u>Organisations-</u> und <u>System-</u> ebene	Ja	Ja	Ja	Ja	Nicht vorgese-hen	Keine	Gut zugänglich

Kriterium:	Anwendungsbereich ⁵³	Gesellschaftliche Ebene der Wirkung	Erfasste Wirkungen		Eignung für:		Aggregation der Daten möglich	Besondere Qualifikationen ⁵⁵	Zugänglichkeit ⁵⁶
			Auch indirekte ⁵⁷	Auch unerwartete	Inter- netes M&E ⁵⁸	Einzelne Studien ⁵⁹ die			
Verfahren:									
Programmiertes integriertes Planungs-, Monitoring- und Evaluationssystem (PrME) von InWEnt	Insbesondere <i>Capacity Building</i> , dabei alle Sektoren und alle Maßnahmentypen, da PME-System	<u>Organisationsebene</u> (Mittler, wenn es sich um <i>Capacity Building</i> handelt), indirekt auch Bevölkerungsebene	Ja (bei Abschluss-evaluierungen)	Ja (bei Abschluss-evaluierungen)	Ja	Nein	Ja, durch gleiche Formate und Anforderungen bei Planung, Monitoring und Evaluierung	Keine	Gut, kann von Webseite von InWEnt heruntergeladen werden
Real Time Evaluation (RTE)	Nothilfe (auf der Mikroebene), insbesondere angewendet für Zwischenevaluierungen	<u>Bevölkerungsebene</u>	eher Nein	eher Ja	Ja	Ja	Nein, Kernfragen können zwischen den Evaluierungen stark variieren	Praktische Nothilfe-Erfahrung erforderlich	Eingeschränkt, vor allem graue Literatur
Verbleibsstudien (<i>Tracker Studies</i>)	Berufliche Bildung, Beschäftigungsförderung, Stipendienprogramme, Hochschulbildung (Mikro/Meso/Makro), <i>Capacity Building</i> in diesen Bereichen	<u>Bevölkerungsebene</u>	Ja	eher Nein	Ja	Ja	Ja	Erstellung und Auswertung von Fragebögen	Leicht eingeschränkt (Publikationen vergriffen)
Wirkungsorientiertes Monitoring (WoM) der GTZ	Geeignet für im bilateralen Auftrag durchgeführte Entwicklungsmaßnahmen; anwendbar auch als Instrument der Steuerung, Rechenschaftslegung und des Wissensmanagements im <i>Multi-Donor</i> -Konzept; je nach der gewählten Interventions-ebene auf der Mikro-, Meso- und Makroebene von Programmen	<u>Bevölkerungsebene</u> , <u>Organisationsebene</u> und <u>Systemebene</u>	Ja	Ja	Ja	Ja	Abhängig von gewählten Indikatoren (quali-quantitativ)	Intensive Befassung mit Thema erforderlich; Erfahrungswissen für Anwendung und Adaption der Methode entscheidend	Eingeschränkt (als GTZ-Publikation vorhanden; Erfahrungswissen z. T. dezentralisiert)

Literatur

Anmerkung: Das Literaturverzeichnis enthält nicht die in Kapitel 2.2 als zusätzliche oder weiterführende Literatur angegebene Literatur.

- AGKED/Misereor (1992): Evaluierung in der kirchlichen Entwicklungsarbeit. Ein Arbeitsbuch. Aachen: Misereor Dialog.
- Anderson, Mary B. (1999): Do No Harm : How Aid can support Peace – or War. Boulder/ London: Lynne Rienner.
- Anderson, Mary B. (Hrsg.) (2000): Options for Aid in Conflict: Lessons from Field Experience. Cambridge, MA: Collaborative for Development Action.
- Anderson, Mary B. und Olson, Lara (2003): Confronting war: critical lessons for peace practitioners. Cambridge, MA: Collaborative for Development Action.
- Appelsmeyer, Heide; Kochinka, Alexander und Straub, Jürgen (1997): Qualitative Methoden. In: Straub, Jürgen; Kempf, Wilhelm und Werbik, Hans (Hrsg.): Psychologie: Eine Einführung. München: Deutscher Taschenbuch Verlag.
- Arbeitskreis „Evaluation von Entwicklungspolitik“ in der DeGEval (Hg.) (2005): Lernen aus Evaluationen. Leitfaden für Geberinstitutionen und GutachterInnen. AG Lernen aus Evaluationen, Arbeitspapiere 2.
- Berg, C. et al. (1998): NGO-Based Participatory Impact Monitoring of an Integrated Rural Development Project in Holalkere Taluk, Karnataka State, India. Berlin: SLE (Schriftenreihe des Seminars für Ländliche Entwicklung Nr. S180).
- Berghof Research Center for Constructive Conflict Management (2005): Berghof Handbook for Conflict Transformation. Dialogue Series No. 4 New Trends in PCIA. Berlin.
- Beywl, Wolfgang (2006): Demokratie braucht wirkungsorientierte Evaluation – Entwicklungspfade im Kontext der Kinder- und Jugendhilfe. In: Deutsches Jugendinstitut (Hg.): Wirkungsevaluation in der Kinder- und Jugendhilfe. München: Deutsches Jugendinstitut.
- Borrmann, Axel und Stockmann, Reinhard (2009): Systemprüfung der Evaluierung in der deutschen Entwicklungszusammenarbeit, Forschungsauftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung (im Erscheinen).
- BMZ (Hrsg.) (2000): Langfristige Wirkungen deutscher Entwicklungszusammenarbeit und ihre Erfolgsbedingungen. Eine Ex-post-Evaluierung von 32 abgeschlossenen Projekten, BMZ Spezial Nr. 019. Bonn.
- BMZ (2003): Deutsche Entwicklungszusammenarbeit. Kurzfassung der Querschnittsanalyse "Monitoringinstrumente in der deutschen EZ". Bonn
- BMZ (Hrsg.) (2005): Ex-Post Evaluation of Sustainability of Regional Rural Development Programmes (RRD). Synthesis Report of Four Country Case Studies. Bonn.
- BMZ (Hrsg.) (2006): Evaluierungskriterien für die deutsche bilaterale Zusammenarbeit. Eine Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen. Bonn.
- Briones, R. et al. (2004): Impact Pathway analysis for research planning: The case of aquatic resources research in the WorldFish Centre. In: NAGA, WorldFish Center Quarterly Vol.27 No.3 & 4 Jul-Dec 2004.

- Bush, Kenneth (1998): A Measure of Peace: Peace and Conflict Impact Assessment of Development Projects in Conflict Zones (Working Paper No.1). The Peace-building and Reconstruction Programme Initiative and The Evaluation Unit, IDRC.
- Caspari, Alexandra (2004): Evaluation der Nachhaltigkeit von Entwicklungszusammenarbeit. Zur Notwendigkeit angemessener Konzepte und Methoden. Wiesbaden: VS-Verlag.
- Caspari, Alexandra und Barbu, Ragnhild (2008): Wirkungsevaluierungen: Zum Stand der internationalen Diskussion und dessen Relevanz für Evaluierungen der deutschen Entwicklungszusammenarbeit. Evaluation Working Papers. Bonn: BMZ.
- Center for Global Development (2006): When will we ever learn? Improving lives through impact evaluation. Report of the Evaluation Gap Working Group. Washington D.C.
- DAC (2002): Glossary of Key Terms in Evaluation and Results Based Management.
- DAC Evaluation Network (2006): DAC Evaluation Quality Standards (for test phase application).
- DAC Working Party on Aid Evaluation (2001): Results Based Management in the Development Cooperation Agencies: A Review of Experience.
- Davies, R. und Dart, J. J. (2005): The 'Most Significant Change' (MSC) Technique. A Guide to Its Use. Cambridge: Self-published.
- DeGEval (2002): Standards für Evaluation (www.degeval.de/calimero/tools/proxy.php?id=70).
- Diekmann, Andreas (2005): Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen. Reinbek: Rowohlt.
- DFID (2004): Introducing Sustainable Livelihoods. A Trainers' Guide.
- Dolzer, Hermann; Dütting, Martin; Galinski, Doris; Meyer, Lutz R. und Rottländer, Peter (1998): Wirkungen und Nebenwirkungen. Ein Beitrag von Misereor zur Diskussion über Wirkungsverständnis und Wirkungserfassung in der Entwicklungszusammenarbeit. Aachen: Misereor.
- Douthwaite, B.; Schulz, S. und Olanrewaju, A. (CGIAR) (2002): Impact Pathway Evaluation: An Approach for Achieving and Attributing Impact in African Agriculture.
- Duflo, Esther; Glennerster, Rachel und Kremer, Michael (2008): Using Randomization in Development Economics Research: A Toolkit. Chapter 61 in T. Paul Schultz and John Strauss (eds.) Handbook of Development Economics, vol. 4, pp 3895-3962. Elsevier: North-Holland.
- Duflo, Esther und Kremer, Michael (2005): Use of Randomization in the Evaluation of Development Effectiveness. In Pitman, George, Feinstein, Osvaldo und Ingram, Gregory (eds.) Evaluating Development Effectiveness. New Brunswick, NJ: Transaction Publishers 2005, S. 205-232.
- Earl, Sarah et al. (2001): Outcome Mapping. Building Learning and Reflection into Development Programs. International Development Research Centre (IDRC).
- Earle, L. (2004): Introduction, In: Earle, L. (Hrsg.) Creativity and Constraint. Grassroots Monitoring and Evaluation and the International Aid Arena. NGO Management & Policy Series 18. Oxford: INTRAC.

- Engelhardt, A. et al. (2003): Impact Assessment tools of DFID's Renewable Natural Resources Research Strategy.
- Englert, Annette (2006): Wirkungsmessung zwischen Anspruch und Wirklichkeit. Alltägliche Herausforderungen bei der Durchführung von Evaluierungen. Diskussionspapier auf der AGEE-Fachtagung zur Evaluationspraxis in der EZ. In: *Entwicklungsethnologie* 15 (1/2).
- European Evaluation Society (2007): The importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions. (EES Statement, December 2007).
- Faust, Jörg (2009): Zuverlässige Wirkungsbelege, in: *E + Z*, Nr.1/2009, S. 14-17.
- Friedrich-Ebert-Stiftung (2007): Peace and Conflict Impact Assessment Methodical Guidelines. Bonn.
- Glewwe, Paul; Kremer, Michael und Moulin, Sylvie (2004): Many Children Left Behind? Textbooks and Test Scores in Kenya. In: *American Journal: Applied Economics*, 1(1): 112-135.
- Germann, D. und Gohl, E. (1996): Participatory Impact Monitoring. Booklet 1: Group-Based Impact Monitoring. Booklet 2: NGO-Based Impact Monitoring. Booklet 3: Application Examples. Booklet 4: The Concept of Participatory Impact Monitoring. Hrsg.: GATE/GTZ und FAKT. Wiesbaden: Vieweg.
- Gohl, Eberhard (2000): Prüfen und Lernen. Praxisorientierte Handreichung zur Wirkungsbeobachtung und Evaluation. Bonn: VENRO.
- Gohl, Eberhard (2007): NGO-IDEAs. Bericht über die Wirkungen von Spar- und Kreditprogrammen süd-indischer NRO. Bensheim (Karl-Kübel-Stiftung).
- Graaff, Jan de (Hrsg.) (2007): Monitoring and evaluation of soil conservation and watershed development projects. Enfield, NH [etc.]: Science Publishers Publication.
- GTZ (2004): Wirkungsorientiertes Monitoring. Leitfaden für Vorhaben der Technischen Zusammenarbeit. GTZ Stabsstelle 04 Unternehmensentwicklung, OE 042 Interne Evaluierung.
- GTZ (2007): Peace and Conflict Assessment (PCA) : Ein methodischer Rahmen zur konflikt- und friedensbezogenen Ausrichtung von EZ-Maßnahmen. Eschborn.
- Hallam, Alistair (1998): Good Practice Review Evaluating Humanitarian Assistance Programmes in Complex Emergencies (Hrsg. Humanitarian Practice Network, 1998).
- Hancock, J.; Proctor, F. und Csaki, C. (2003): Scaling-up the Impact of Good Practices in Rural Development. A Working Paper to Support Implementation of the World Bank's Rural Development Strategy. The World Bank Agriculture & Rural Development Department Report No. 26031.
- Herweg, Karl (2005): IMA – Main Messages and Experiences. Bern.
- Herweg, Karl und Steiner, Kurt (2002): Impact Monitoring & Assessment. Instruments for use in rural development projects with a focus on sustainable land management. Volume 1: Procedure. Volume 2: Toolbox. Bern/Eschborn: CDE/GTZ et al.
- ICCO (2000): Building Bridges in PME. Guidelines for good practice in the planning, Monitoring and Evaluation of Community-Based Development Projects implemented by Southern NGOs with support from European Ecumenical Agencies. Zeist, www.icco.nl.

- IDD et al. (2006): Evaluation of General Budget Support: Synthesis Report. Birmingham.
- Joint Marrakech Memorandum (2004), Annex 1: Promoting a harmonized approach to managing for development results: Core principles.
- Karl Kübel Foundation for Child and Family (2007): The Impact Tool Box. Version 1.1, 2007. Tool Box for Impact Monitoring and Evaluation of Savings & Credit Programme. Cochin, <http://www.ngo-ideas.net/toolbox.html?> (1.12. 2008).
- Kremer, Michael (2003). Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons, American Economic Review Papers and Proceedings 93(2): 102-115.
- Lange, R. (2001): Tracer Studies. Working instrument for a new orientation of vocational training. Stuttgart: AGKED/FAKT (www.fakt-consult.de).
- Leeuw, Frans und Vaessen, Jos (2009): Impact Evaluations and Development. NONIE Guidance on Impact Evaluation. Draft Version for Discussion, April 2009.
- Lunch, Chris (2007): The Most Significant Change: using participatory video for monitoring and evaluation. In: Participatory Learning and Action, Volume 56, Number 1, June 2007, S. 28-32(5).
- Mosley, P. (1987): Overseas Aid: its Defence and Reform. Brighton: Wheatsheaf.
- Neubert, Susanne (2004): Wirkungsanalysen der entwicklungspolitischen Zusammenarbeit sind machbar. DIE. Analyse und Stellungnahmen, Nr. 4. Bonn.
- Neubert, Susanne (2004a): Impact Analysis of Development Cooperation is Feasible. Briefing Paper 4/2004, DIE, Bonn.
- Neubert, Susanne (2004b): Akteurszentrierte Wirkungsanalyse – konzeptionelle Überlegungen und Scaling up für die Verwaltungszusammenarbeit. In: Simon, Klaus (Hrsg.) (2004): Verwaltungen, Experten und Bürger im Reformprozess – Wirkungen und Evaluierung von Verwaltungszusammenarbeit mit Entwicklungsländern. Baden-Baden: Nomos Verlag.
- Neubert, Susanne (2007): Zur Wirkungserfassung: Ein Überblick über die Methoden aus Praxis und Wissenschaft. Tagungsband zum 10. Entwicklungsökonomischen Seminar: Zur Wirkungsdebatte in der Entwicklungszusammenarbeit der KfW am 17-20. Januar 2007 in Erfurt.
- OECD (2005): Paris Declaration on Aid Effectiveness. Ownership, Harmonisation, Alignment, Results and Mutual Accountability. Paris.
- Paffenholz, Thania (2005): Third-generation PCIA : Introducing the Aid for Peace Approach. (www.berghof-handbook.net/uploads/download/dialogue4_paffenholz.pdf).
- Paffenholz, Thania und Reyhler, Luc (2007): Aid for Peace : A Guide to Planning and Evaluation for Conflict Zones. Baden-Baden: Nomos Verlag.
- Patton, Michael Quinn (2002): Qualitative research & evaluation methods. Thousand Oaks/London/New Delhi: Sage.
- Post, U. und Warning, C. (2007): Begründetes Unbehagen. In: E+Z, 48(2), 72–74, Frankfurt .
- Ravallion, Martin (2001): The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation, World Bank Economic Review 15(1), 115-14.

- Ravallion, Martin (2008): Evaluating Anti-Poverty Programs. Chapter 59 in T. Paul Schultz and John Strauss (eds.) *Handbook of Development Economics*, vol. 4, pp 3787-3846. Elsevier: North-Holland.
- Rogers, P. J. et al. (2000). Impact Pathway evaluation: practice, promise, and problems. In: Rogers, P. J. et al.: *Impact Pathway in Evaluation. Challenges and Opportunities*. New Directions for Evaluation 87.
- Rossi, Peter H.; Lipsey, Mark und Freeman, Howard E. (2004): *Evaluation: a systematic approach*. Thousand Oaks/London/New Delhi: Sage.
- Santha, E. K. (2007): *Small Credit: Big Impact. A Report on the Savings & Credit Programme of NGO-IDEAs Partner Organisations*. Coimbatore (Karl Kübel Foundation for Child and Family).
- Schnell, R.; Hill, P.B. und Esser, E. (2008): *Methoden der empirischen Sozialforschung*. München: Oldenbourg Wissenschaftlicher Verlag.
- Schubert, B.; Agrawal, R. C.; Böttcher, D. u. a. (1984): *Die Nachhaltigkeit der Wirkungen von Agrarprojekten. Eine Querschnittsanalyse von Projekten der deutschen Technischen Zusammenarbeit*. Forschungsberichte des Bundesministeriums für wirtschaftliche Zusammenarbeit Nr. 64. München: Weltforum.
- Schürmann, A. (2002): *Participatory Impact Monitoring of Self-Help Groups and Watersheds. A Users' Handbook*. Bonn: Welthungerhilfe.
- Sigsgaard, P. (2004): *Doing away with Predetermined Indicators: Monitoring Using the Most Significant Changes Approach*. In: Earle, L. (Hrsg.): *Creativity and Constraint. Grassroots Monitoring and Evaluation and the International Aid Arena*. NGO Management & Policy Series 18. Oxford: INTRAC.
- Stockmann, Reinhard (2006a): *Evaluation und Qualitätsentwicklung*. Münster: Waxmann.
- Stockmann, Reinhard (2006b): *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*. Band 1 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann, dritte überarbeitete und aktualisierte Auflage.
- Stockmann, Reinhard (Hrsg.) (2007): *Handbuch zur Evaluation. Eine praktische Handlungsanleitung*. Band 6 der Reihe Sozialwissenschaftliche Evaluationsforschung. Münster: Waxmann.
- Stockmann, Reinhard (2008): *Evaluation and Quality Development. Principles of Impact-Based Quality Management*. Frankfurt am Main: Peter Lang.
- Strele, M. et al. (2006a): *Linking Programmes and Poor People's Interests to Politics Experiences from Cambodia*. FAO Livelihood Support Programme. März 2006.
- Strele, M. et al. (2006b): *Participatory Impact Monitoring for Sustainable Livelihoods*. Paper at EASY ECO Conference. Saarbrücken.
- The Advisory Group on Civil Society & Aid Effectiveness (2008): *Synthesis of Findings and Recommendations*. Ottawa.
- The North-South Civil Society Dialogue on Aid Effectiveness (2007): *Nairobi Declaration – Civil Society Resolution on the Effectiveness of Development Cooperation and the Paris Declaration*, in: The Advisory Group on Civil Society and Aid Effectiveness (2007): *A North-South Civil Society Dialogue (Report of the Nairobi Meeting November 15-16 2007)*, nur im Internet veröffentlichtes Dokument: http://www.concordeurope.org/Files/media/internetdocuments/ENG/3_Topics/Topics/20_CSO_effectiveness/Final-Report-of-the-Nairobi-N-S-CSO-Dialogue.doc.

- United Nations Volunteers (Hrsg.) (2008): A Participatory Methodology for Assessing the Contribution of Volunteering to Development. A Handbook for Volunteers and Programme Officers. Bonn.
- UN Joint Inspection Unit (2004): Implementation of Results-Based Management in the United Nations Organizations.
- VENRO (2007): Mein Wort zählt. Mikrokredite: Kleines Kapital – große Wirkung. Frankfurt/Main.
- Vester, F. (2002): Die Kunst vernetzt zu denken. München: Deutscher Taschenbuch Verlag.
- Weingärtner, L. et al. (2005): Poverty and Food Security Monitoring in Cambodia. Berlin: SLE.
- Welthungerhilfe (2008a): Fachkonzept Wirkungsorientierte Evaluation der Auslandsarbeit der Welthungerhilfe. Bonn.
- Welthungerhilfe (2008b): Wirkungen erzielen – Leitfaden zur Wirkungsorientierung in den Projekten und Programmen der Welthungerhilfe. Bonn.
- Willetts, Juliet und Crawford, Paul (2007): The most significant lessons about the Most Significant Change technique. In: Development in Practice, June 2007, Vol. 17, Issue 3, S. 367-379.
- Wolff, Jürgen H. (2005): Entwicklungshilfe: Ein hilfreiches Gewerbe? Münster: Lit-Verlag.
- Wouters, Shirley et al. (2003): Development of a Peace and Conflict Impact Assessment for Communities in the South Caucasus. Berlin: Seminar für Ländliche Entwicklung (Schriftenreihe des SLE Nr. S206).
- Zintl (2006): Von Wirkungen und Nebenwirkungen – Evaluierung nach der Millenniumserklärung und der Pariser Erklärung. In: EINS Entwicklungspolitik, 9-2006.
- Zürcher, Christoph; Köhler, Jan und Böhnke, Jan (2007): Assessing the Impact of Development Cooperation in North East Afghanistan, Interim Report, Evaluation Reports 028. Bonn: BMZ.
- Zürcher, Christoph und Köhler, Jan (2007): Assessing the Impact of Development Cooperation in North East Afghanistan: Methods and Approaches; Evaluation Working Papers. Bonn: BMZ (*auf Anfrage erhältlich*).

Glossar

*Begriffe werden in der Entwicklungszusammenarbeit sehr unterschiedlich verwendet. In diesem Glossar wird dargestellt, wie diese Publikation Begriffe verwendet, z. T. mit Verweis darauf, dass es auch andere Verwendungen gibt.**

Aktivitäten: Gesamtheit von Aktionen (Maßnahmen), die im Rahmen einer Entwicklungsmaßnahme durchgeführt werden.

Ansatz: Eine (wissenschaftliche) Herangehensweise an den Forschungs-/Evaluierungsgegenstand, die Elemente aus Theorie, Methode und Forschungs-/Evaluationstechnik miteinander kombiniert. CEval spricht von einem „theoretischen Gesamtkonzept zur Evaluation“.

Design: Die gesamte Gestaltung des Vorgehens zur Beantwortung einer (wissenschaftlichen) Fragestellung (z. B. experimentelles oder nicht-experimentelles Design). Dazu gehört die Auswahl von Verfahren und Instrumenten und der Zeitplan unter den gegebenen Restriktionen (Geld, Zeit, Personal, Raum).

Entwicklungsmaßnahme: Jegliche abgrenzbare Maßnahme (Projekt, Programm, Politik usw.) zur Erzielung einer entwicklungspolitischen Wirkung.

Evaluation/Evaluierung: Nach OECD/DAC die systematische und objektive Bewertung eines laufenden oder abgeschlossenen Projekts, Programms oder einer Politik, seines Designs, seiner Durchführung und Ergebnisse (*results*). Nach Bewyl (2004) bezeichnet Evaluation die Summe systematischer Untersuchungen, die empirische, d. h. erfahrungsbasierte Informationen bereitstellen, so dass es möglich wird, den Wert (Güte) eines Vorhabens einzuschätzen.

Impact: Nach OECD/DAC positive und negative, primäre und sekundäre, häufig längerfristige Effekte, ausgelöst durch eine Entwicklungsmaßnahme, direkt oder indirekt, beabsichtigt oder unbeabsichtigt.

Indikatoren: Indikatoren sind Anzeiger (Messgrößen), die es erlauben, begründet auf nicht unmittelbar wahrnehmbare Sachverhalte zu schließen. Nach OECD/DAC ist ein Indikator ein/e quantitative/r oder qualitative/r Faktor oder Variable, der/die eine einfaches und verlässliches Mittel bietet, Erfolg zu messen, die Veränderungen widerzuspiegeln, die mit einer Intervention verbunden sind, oder die helfen, die Leistung eines entwicklungspolitischen Akteurs zu bewerten.

Input: Finanzielle, personelle, materielle Beiträge und Mittel, die im Rahmen einer Entwicklungsmaßnahme von verschiedenen Beteiligten und durch verschiedene Instrumente bereit gestellt werden.

Instrument: Ein Werkzeug oder Hilfsmittel zur Erfüllung eines ganz bestimmten Zwecks, z. B. zur Datenerhebung oder zur Datenauswertung.

Konservierende Wirkungen: Die Aufrechterhaltung eines positiven Zustands durch eine Entwicklungsmaßnahme (im Gegensatz zu einer bewirkten „Veränderung“), vor allem in einem sich verschlechternden Umfeld, z. B. im ökologischen Bereich.

* Wo nicht anders ausgewiesen, richten sich die Definitionen nach Bewyl (2004, <http://www.univention.org/glossar/index.php>) und nach BMZ (2006, http://www.bmz.de/de/zen_trales_downloadarchiv/erfolg_und_kontrolle/evaluierungskriterien.pdf) (Wirkungskette).

Kontrafaktisch: Gedanklicher Vergleich der realen Entwicklung (mit Entwicklungsmaßnahme) mit der imaginären Entwicklung (ohne Entwicklungsmaßnahme): Was wäre ohne die Intervention passiert?

Kontextorientiert: Auf die nationale oder lokale Lebenswirklichkeit der Menschen bezogen. Von der Situation und Veränderung vor Ort ausgehend (im Gegensatz zu „hypothesengeleitet“).

Kontrollgruppe: Eine Gruppe, die im Rahmen eines Experimentaldesigns eine Basis für den Vergleich mit der Untersuchungsgruppe liefert. Die Kontrollgruppe wird dem Vorhaben nicht ausgesetzt. Damit ist ein Vergleich mit der Untersuchungsgruppe möglich, die von dem Vorhaben betroffen ist.

Konzept: Ein vereinfachtes Bild eines Ausschnitts der Realität, es kann als Vorstufe einer Theorie dienen.

Methode: Systematische, d. h. nach festgelegten Regeln ablaufende Vorgehensweise der Sozialforschung vor allem zur Datengewinnung und -auswertung.

Mikro-Makro-Paradoxon: Der (scheinbare) Gegensatz, der sich daraus ergibt, dass die überwiegende Zahl von Entwicklungsmaßnahmen als positiv wirksam evaluiert werden und gleichzeitig in ländervergleichenden ökonometrischen Untersuchungen kein robuster Zusammenhang zwischen Entwicklungshilfe(inputs) und Entwicklung (auf der Makroebene) ermittelt werden kann.

Mikro-, Meso- und Makroebene: Auf der Mikroebene (lokalen Ebene) werden Individuen, Haushalte oder Gruppen/Gemeinden betrachtet (Hauptwirkungsebene), auf der Makroebene (nationalen Ebene) die Gesellschaft (z. B. innerhalb eines Staates), im politischen Sinne jedoch die nationalen Organe und Gesetzgebungen (Regierung, Hauptentscheidungsebene). Auf der dazwischen liegenden Mesoebene (regionalen Ebene) werden Administrationen (z. B. Landesbehörden), nichtstaatliche und private Dienstleister, Verbände oder Netzwerke betrachtet (Hauptumsetzungsebene).

Mit-Ohne-Vergleich: s. $\Rightarrow$ Kontrollgruppe

Mittler: Akteure zwischen einem Vorhaben und seiner Zielgruppe (z. B. Behörden, Organisationen, Berater), über die entwicklungspolitische Wirkungen vermittelt werden.

Modell: Eine gedanklich konstruierte, für einen bestimmten Zweck entworfene, bewusst vereinfachende Nachbildung grundlegender Merkmale eines Sachverhalts oder Vorgangs.

Monitoring: Die unmittelbare, systematische und regelmäßige Erfassung eines Vorgangs oder Prozesses mit Hilfe von Beobachtungssystemen. Nach OECD/DAC hat Monitoring „a continuing function that uses systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.“ Monitoring bezieht sich in Vorhaben der EZ nicht in erster Linie auf Wirkungen, sondern vor allem auch auf Aktivitäten und Ergebnisse. Wirkungsmonitoring stellt also einen Teilaspekt des Projekt- oder Programmmonitoring dar.

Nutzen: Beteiligte profitieren von der Nutzung der $\Rightarrow$ Outputs einer Entwicklungsmaßnahme. Der Nutzen ist Teil des $\Rightarrow$ *outcome*.

Nutzung: Beteiligte nutzen die erstellten Produkte, Güter, Dienstleistungen und Institutionen/Regelwerke (den $\Rightarrow$ Output). Im Verständnis des BMZ-Evaluierungsreferats ist die Nutzung Teil des $\Rightarrow$ *outcome*.

Outcome: (Direkte) Wirkungen der Leistungen einer Intervention, die Zuordnung ist in der Regel möglich.

Output: Produkte, Güter, Dienstleistungen und Institutionen/Regelwerke, die durch eine Entwicklungsmaßnahme erstellt werden. Nach OECD/DAC können Outputs „Veränderungen beinhalten, die von einer Intervention ausgehen und die relevant für die Erreichung des *outcome* sind“ (etwa unmittelbare Umweltauswirkungen im Produktionsprozess).

Real Time Evaluation: „A real time evaluation is a timely, rapid and interactive peer review in fast evolving humanitarian operations undertaken at an early stage. Its objective is to gauge the effectiveness and impact of a given response, and to ensure that its findings are used as an immediate catalyst for organizational and operational changes“ (UNHCR, EPAU 2002).

Sättigung(sgrad): Bezieht sich auf ein im Bereich der qualitativen Sozialforschung formuliertes Ziel (oder Kriterium), wonach der untersuchte Gegenstandsbereich (das untersuchte Phänomen) so weit erschlossen werden soll, dass durch neue Daten und weitere Auswertungen keine neuen Erkenntnisse mehr erwartet werden können. Ähnliche Begriffe sind empirische Sättigung, theoretische Sättigung.

Sector Wide Approach (SWAp): Ein programmbasierter sektoraler Ansatz im Sinne eines „way of engaging in development cooperation based on the principle of coordinated support for a locally owned programme of development“ (OECD DAC).

Selbstevaluierung: Systematische und datenbasierte Bewertung der eigenen Tätigkeit. Eine Selbstevaluierung wird mit dem Ziel durchgeführt, Qualität (und ggf. Wirkungen) nachzuweisen und zu verbessern. Die Akteure befinden sich somit in einer Doppelrolle (des Bewerteten und Bewertenden).

Verfahren: Ein geregelter, in einzelne Schritte zerlegbarer, transparenter und wiederholbarer Ablauf, um ein bestimmtes Ziel zu erreichen, in dem verschiedene theoretische Konzepte, Modelle und Instrumente enthalten sein können. In einem Verfahren werden bestimmte $\Rightarrow$ Methoden konkret angewendet bzw. umgesetzt.

Vorhaben: s. Entwicklungsmaßnahme

Vorher-Nachher-Vergleich: Vergleich der Situation einer Entwicklungsmaßnahme (nach einer gewissen Laufzeit, am oder nach ihrem Ende) mit der Situation zu einem früheren Zeitpunkt.

Wirkungen: Nach BMZ (2006) auf der Ebene des $\Rightarrow$ *outcome* die wahrscheinlichen oder tatsächlich erreichten kurz- und mittelfristigen Veränderungen in Folge der Nutzung der Leistungen einer Intervention, auf der Ebene des $\Rightarrow$ *impact* (s. dort).

Wirkungsanalyse/Wirkungsevaluierung: In einer Wirkungsanalyse werden entwicklungspolitische Wirkungen sozialwissenschaftlich nachgewiesen bzw. plausibilisiert. In einer Wirkungsevaluierung werden die Wirkungen zusätzlich bewertet.

Wirkungshypothese: Begründete Vermutung über Wirkungen eines Vorhabens.

Wirkungsmodell: s. $\Rightarrow$ Modell

Wirkungsmonitoring: Auf Wirkungsindikatoren bezogenes $\Rightarrow$ Monitoring.

Wirkungsorientiertes Monitoring: Wirkungsorientiertes Monitoring umfasst das gesamte Monitoring, das auf Wirkung ausgerichtet wird und Wirkungsmonitoring umfasst.

Zielgruppe: Die Bevölkerungsgruppen, denen eine Entwicklungsmaßnahme letztendlich zugute kommen soll („Endbegünstigte“).

Mitglieder der AG „Wirkungsanalyse“

AFC Consultants International (Matthias Webendörfer u. a.)
 AGEG Consultants eG (Theo Mutter)
 AGEH – Arbeitsgemeinschaft für Entwicklungshilfe (Petra Feil, Beraterin)
 BfdW – Brot für die Welt (Gerhild Gürtler u. a.)
 BMZ – Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung
 CBM – Christoffel Blindenmission (Thomas Hochgesang, Stefan Dofel)
 CEVAL – Centrum für Evaluation (Peter Maats, Nicolás Reade, Stefan Silvestrini)
 DIE – Deutsches Institut für Entwicklungspolitik (Dr. Susanne Neubert u. a.)
 EED – Evangelischer Entwicklungsdienst (Kirsten Vorwerk u. a.)
 FAKT Consult GmbH (Rudolf Holtkamp, u. a.)
 GTZ – Gesellschaft für Technische Zusammenarbeit GmbH (Thomas Grammig u. a.)
 HBS – Heinrich-Böll-Stiftung (Christiane Dilger)
 Impact Plus – (Dr. Eberhard Gohl, Verena Brenner)
 InWent – Internationale Weiterbildung und Entwicklung gGmbH (Rita Walraf u. a.)
 KfW – Kreditanstalt für Wiederaufbau (Holger Papenfuß)
 Milango GmbH (Sigfrid Schröder-Breitschuh)
 Misereor (Dr. Hermann Dolzer u. a.)
 terres des hommes (Klaus Müller-Reimann u. a.)
 United Nations Volunteers (Caspar Merkle)
 Universität Köln (Martin Quack)
 Welthungerhilfe (Dr. Dirk Guenther, Ute Rössing)
 Freie Gutachter/-innen:
 Christian Berg (Koordinator der AG), Bernward Causemann, Boris Keller, Maria
 Lehmann, Regina Müller, Dr. Bettina Schmidt, Dr. Karin Stahl

Autorinnen und Autoren

Christian Berg arbeitet als freier entwicklungspolitischer Gutachter und Berater, vorwiegend im Bereich Monitoring und Evaluierung, ist freier Mitarbeiter, Teamleiter und Trainer am Seminar für Ländliche Entwicklung (SLE) der Humboldt-Universität zu Berlin und Gesellschafter der Beratungsfirma comit GmbH, Berlin. E-Mail: c.berg@comit-berlin.net

Karola Block ist freie Organisationsberaterin in der Entwicklungszusammenarbeit mit den Schwerpunkten Change Management, Wirkungsbeobachtung, Konfliktbearbeitung, Intensive Beratung, Mitkonzeption und Training zu PriME für InWEnt. E-Mail: karola.block@t-online.de

Bernward Causemann ist freier entwicklungspolitischer Berater mit Schwerpunkt auf Organisationsentwicklung, Wirkungsbeobachtung, Prozessbegleitung und Coaching von Beratern. Er war bis 2007 Leiter der Managementberatung von FAKT gGmbH. E-Mail: bc@causemann.org

Claudia Conrad ist Politologin mit langjähriger Erfahrung in der Internationalen Zusammenarbeit und heute als Organisations- und Kooperationsmanagementberaterin und Evaluatorin tätig. Sie ist Partnerin von KEK-CDC Consultants in Zürich. E-Mail: conrad@kek.ch

Stefan Dofel arbeitet gegenwärtig als Programmleiter im Auftrag der CBM Christoffel-Blindenmission für die Regierung von Malawi und ist als entwicklungspolitischer Gutachter insbesondere in den Bereichen Evaluation, Organisationsentwicklung und Management von Veränderungsprozessen tätig. E-mail: stefandofel@gmx.de

Dr. Achim Engelhardt arbeitet als M&E-Consultant assoziiert mit der Lotus M&E Group. Bis 2006 arbeitete er für das DFID M&E Resource Centre „Parc“ in Birmingham im Bereich „Donor Harmonisation“, „Policy Evaluation“, „Aid Effectiveness“ und „Results-Based Management“. Er ist Gründungsmitglied der Arbeitsgruppe für Programmevaluierung in der Entwicklungszusammenarbeit der Schweizer Evaluationsgesellschaft. E-Mail: achim@lotus-group.org

Markus Fiebiger ist seit 2000 in unterschiedlichen Rollen in der internationalen Entwicklungszusammenarbeit tätig, Arbeitsschwerpunkte sind Projektplanung in der ländlichen Entwicklung, nachhaltiges Ressourcenmanagement und Landnutzung, ökologische Risikobewertung, Landschafts- und Agrarökologie sowie partizipative Planungs- und Monitoringmethoden. E-Mail: markus.fiebiger@gmail.com

Prof. Dr. Markus Frölich lehrt und forscht an der Universität Mannheim insbesondere im Bereich mikroökonomische Methoden zur Evaluation sowie angewandte Evaluationsstudien und ist Direktor des Forschungsschwerpunkts Beschäftigung und Entwicklung des Instituts zur Zukunft der Arbeit (IZA) und der Weltbank. E-Mail: froelich@uni-mannheim.de

Dr. Eberhard Gohl ist als Inhaber von Impact Plus spezialisiert auf Wirkungsanalysen sowie die Stärkung von wirkungsorientierter Steuerung von Projekten und Organisationen, bei FAKT koordiniert er den Bereich „Wirkungen“. E-Mail: gohl@impact-plus.de

Claudia Gottmann ist Projektmitarbeiterin im Querschnittsvorhaben Armutsbekämpfung der GTZ und dort verantwortlich für den Themenbereich armutsorientierte Wirkungsbeobachtung. E-Mail: claudia.gottmann@gtz.de

Dr. Dirk Guenther leitet die Stabsstelle Evaluation und Qualitätsmanagement der Welthungerhilfe in Bonn. Zuvor leitete er von 1999 bis 2003 das Regionalbüro der Welthungerhilfe in Port au Prince und anschließend bis 2005 die Regionalgruppe Süd- und Südostasien in Bonn. E-Mail: dirk.guenther@welthungerhilfe.de

Thomas Hochgesang befasst sich in der CBM Christoffel-Blindenmission (Christian Blind Mission) e.V. seit 2005 schwerpunktmäßig mit den Themen Aid Effectiveness und Wirkungserfassung und ist Sprecher der AG Wirkungsbeobachtung des Verbandes Entwicklungspolitik deutscher Nichtregierungsorganisationen (VENRO). E-Mail: thomas.hochgesang@cbm.org

Prof. Dr. Stefan Klonner ist Professor für Außenwirtschaft und Entwicklungsökonomie mit Forschungsschwerpunkten in Armutsmessung und -entwicklung sowie Kreditmärkte in Entwicklungsländern an der Johann-Wolfgang-Goethe-Universität, Frankfurt am Main. E-Mail: klonner@wiwi.uni-frankfurt.de

Ralf Lange ist Geschäftsführer der FAKT GmbH und als Berater in den Bereichen Berufliche Bildung, Beschäftigungsförderung – Arbeitsmarkt und Wiederaufbau in Postkonfliktsituationen tätig. Arbeitsschwerpunkte sind Evaluation sowie die Entwicklung wirkungsorientierter Monitoringsysteme. E-Mail: ralf.lange@fakt-con sult.de

Dr. Matthias Lanzendörfer arbeitet bei Misereor im Bereich Evaluierung und Qualitätsmanagement (EQM). E-Mail: matthias.lanzendoerfer@misereor.de

Caspar Merkle arbeitet seit 2006 beim Freiwilligenprogramm der Vereinten Nation (UNV) im Bereich Monitoring und Evaluierung und war zuvor mehrere Jahre für die Internationale Arbeitsorganisation (ILO) tätig, mit den Arbeitsschwerpunkten Projekt- und Programmevaluierungen sowie Wirkungsanalyse. E-Mail: caspar.merkle@unvolunteers.org

Regina Müller ist Sozialwissenschaftlerin und freie Gutachterin und Beraterin in den Bereichen Bildung und ländliche Entwicklung sowie Mitglied der DeGEval. E-Mail: regina.mueller@t-online.de

Sonja Nelles ist Projektleiterin in der Abteilung Qualität und Evaluierung bei InWEnt, steuert im Rahmen des Qualitätsmanagements interne Lern- und Verbesserungsprozesse und ist zuständig für die konzeptionelle Weiterentwicklung von PRiME. Darüber hinaus führt sie interne Trainings durch und berät die Kolleginnen und Kollegen bei der Anwendung von PRiME. E-Mail: sonja.nelles@inwent.org

Dr. Susanne Neubert ist Wissenschaftlerin am Deutschen Institut für Entwicklungspolitik und dort für die Bereiche „Methoden und Wirkungsevaluierung“ sowie „nachhaltiges Ressourcenmanagement“ zuständig. E-Mail: susanne.neubert@die-gdi.de

Martin Quack ist Politikwissenschaftler mit dem Schwerpunkt Friedens- und Konfliktforschung, Research Fellow am Lehrstuhl für Internationale Politik und Außenpolitik der Universität zu Köln und arbeitet seit 2008 für das Forum Ziviler Friedensdienst in Kosovo. E-Mail: martin_quack@web.de

Nicolà Reade ist wissenschaftliche Mitarbeiterin am Centrum für Evaluation (CEval) an der Universität des Saarlandes, dort zuständig für den Weiterbildungsbe-
reich, und koordiniert die Evaluationsvorhaben im Auftrag der GTZ. Darüber hinaus
wirkt sie an der konzeptionellen Weiterentwicklung des Ansatzes zur Wirkungseva-
luation (RIE) mit und ist als Expertin für die Implementierung von M&E-Systemen
im Einsatz. E-Mail: n.reade@ceval.de

Dr. Bettina Schmidt unterrichtet Diversity Management in International Business
Administration und in Gesundheitsmanagement, leitete verschiedene Programme im
Auftrag staatlicher und nicht-staatlicher Organisationen und ist Vorstandsmitglied
der Arbeitsgemeinschaft Entwicklungsethnologie und der International Society for
Diversity Management. E-Mail: b.schmidt2@t-online.de

Sigfrid Schröder-Breitschuh verantwortet als Berater und leitender Konzeptionist
die Bereiche Monitoring und Evaluierung sowie Management komplexer EZ-
Vorhaben bei milango GmbH. Neben der prozessbegleitenden Beratung von Län-
derbüros, Programmen und Projekten arbeitet er als Moderator, Coach und Trainer.
E-Mail: sigfrid@milango.com

Dr. Arno Sckeyde arbeitet im Bereich „Planung und Entwicklung“ der GTZ und ist
dort u. a. für die Weiterentwicklung des Methodenproduktes „Wirkungsorientiertes
Monitoring“ zuständig. E-Mail: arno.sckeyde@gtz.de

Stefan Silvestrini ist Senior Wissenschaftler und Bereichskoordinator „Entwick-
lungszusammenarbeit und Bildung“ am Centrum für Evaluation (CEval) an der Uni-
versität des Saarlandes, dort verantwortlich für die Qualitätssicherung im Bereich
der Auftragsforschung, und arbeitet an der theoretischen und methodischen Weiter-
entwicklung des CEval-Ansatzes, insbesondere im Hinblick auf dessen Anwendung
im Rahmen von ex-ante Evaluationen mit. E-Mail: s.silvestrini@ceval.de

Dirk Sprenger arbeitet als Personalentwickler mit den Schwerpunkten Training,
Coaching, Moderation und Mediation und beschäftigt sich im Rahmen der Wir-
kungsorientierung primär mit der Frage nach Wirkungen von Transformationsansät-
zen in Konfliktsystemen. E-Mail: dirk.sprenger@kontrair.de

Siegrid Tautz ist geschäftsführende Gesellschafterin des internationalen Beratungs-
unternehmens evaplan GmbH am Universitätsklinikum Heidelberg und fachlich zu-
ständig für Sexuelle und Reproduktive Gesundheit und Rechte, HIV & AIDS,
Gesundheitsförderung, Curriculumentwicklung und Angewandte Forschung. E-Mail:
siegrid.tautz@evaplan.org

Kirsten Vorwerk ist Referentin für Evaluation beim Evangelischen Entwicklungs-
dienst. E-Mail: kirsten.vorwerk@eed.de

Michaela Zintl leitet das Referat Evaluierung der Entwicklungszusammenarbeit;
Außenrevision im Bundesministerium für wirtschaftliche Zusammenarbeit und Ent-
wicklung (BMZ) und war bzw. ist Mitglied verschiedener Arbeitsgruppen des Eva-
luierungsnetzwerkes des Entwicklungsausschusses der OECD (DAC). Bereits in ih-
rer vorherigen Position als Exekutivdirektorin der Interamerikanischen Entwick-
lungsbank hat sie sich u. a. als Vorsitzende des Policy and Evaluation Committee mit
Wirkungs- und Evaluierungsfragen befasst. E-Mail: michaela.zintl@bmz.bund.de