

## Gemeinsame Frühjahrstagung der AG ‚Evaluation in der Entwicklungszusammenarbeit‘ der SEVAL und des AK ‚Evaluation von Entwicklungspolitik‘ in der DeGEval am 27./28.04.2006 an der ETH Zürich

*Peter Maats*

*Centrum für Evaluation (CEval), Universität des Saarlandes*

Der AK Evaluation Entwicklungspolitik führte in diesem Jahr erstmalig seine jährliche Frühjahrstagung gemeinsam mit der Arbeitsgruppe ‚Evaluation in der Entwicklungszusammenarbeit‘ der SEVAL an der ETH in Zürich durch. Diskutiert wurde das Thema ‚Länderprogrammevaluationen‘. Insgesamt nahmen 52 Teilnehmer und Teilnehmerinnen aus Deutschland und der Schweiz an der Tagung teil.

Spätestens mit der Formulierung der ‚Millenium Development Goals‘ (MDGs) im Jahre 2000 hat sich die Weltgemeinschaft für eine intensivere Zusammenarbeit im Bereich der internationalen Entwicklungszusammenarbeit ausgesprochen und erstmalig sektorspezifische Wirkungsvorgaben formuliert, die bis 2015 sowohl erfüllt sein sollen als auch überprüft werden müssen. Als ein direktes Ergebnis dieses internationalen Konsensus begannen viele Geberländer ihre doch häufig wenig verknüpften Einzelprojekte in den Partnerländer zu strategischen Länderprogrammen, die der Erfüllung der MDGs verpflichtet sein sollten, zu bündeln. Verschiedene Ansätze der Wirkungsevaluation auf Projektebene wurden schon auf der Jahrestagung der DeGEval in Essen vorgestellt. Die Frühjahrstagung in Zürich ging thematisch einen Schritt weiter und diskutierte bestehende Ansätze, ganze Programmkonzepte zu evaluieren und deren Versuche, Wirkungen, vor allem vor dem Hintergrund der Erreichung der MDGs, auch auf Länderebene bewerten zu können. Der nächste logische Schritt, dass Länderprogramme der verschiedenen Geber zu sog. Multi-Donorprojekten und -programme zusammengeführt werden und folgerichtig auch gemeinschaftlich in Form von ‚Joint Evaluations‘ bewertet werden müssen, wird Thema der nächsten Jahrestagung in Lüneburg sein. Es sei noch hinzugefügt, dass das Thema ‚Wirkungsanalyse‘ dem Arbeitskreis so wichtig ist, dass auf der letzten Jahrestagung in Essen eine eigene Arbeitsgruppe zu diesem Thema konstituiert wurde, die seitdem schon mehrfach zusammen gekommen ist.

Zur Thematik der Länderprogrammevaluationen konnten folgende Referenten und Referentinnen gewonnen werden:

1. Prof. Dr. Reinhard Stockmann (CEval): Wirkungsanalysen in einer veränderten Welt

2. Peter Maats (CEval): Ein metaanalytischer Blick auf ausgewählte Landesprogrammevaluationen
3. Dr. Pedro Morazán (freier Gutachter, Südwind Institut Siegburg): Themen, Fragen, Methoden vom BMZ – Länderprogrammevaluationen – ein Bericht aus der Werkstatt; Fallbeispiel Honduras
4. Dieter Zürcher (NADEL): Querschnittsanalysen von Länderprogrammevaluationen verschiedener Geber
5. Pierre Walther (Walther Consulting): Impact Stories – ein Platz für das Narrative in Programmevaluationen
6. Claudia Conrad (KEK): Real-time Evaluierungen von Programmen der Humanitären Hilfe
7. Monica Rubiolo (seco): Zur Methodik von Budgethilfeevaluationen – Erste Erfahrungen der seco (Staatssekretariat für Wirtschaft)

Alle Vorträge stehen auf der Homepage des Arbeitskreises<sup>1</sup> zum Download zur Verfügung.

Zu Beginn der Tagung wurden die Teilnehmerinnen und Teilnehmer von Herrn Dieter Zürcher als Gastgeber der Tagung, von Herrn Dr. Thomas Widmer stellvertretend für den Vorstand der SEVAL und durch Herrn Prof. Dr. Reinhard Stockmann als Sprecher der Arbeitskreises Entwicklungspolitik in der DeGEval begrüßt.

Anschließend eröffnete Herr Prof. Dr. Reinhard Stockmann die Tagung mit einem einführenden Vortrag, in dem er das Instrument der Länderprogrammevaluationen in den Kontext von Wirkungsevaluationen stellte. Er betonte, dass die Ziele von entwicklungspolitischen Projekten und Programmen in den letzten Jahren u.a. durch Formulierung der MDGs höher gesteckt wurden und deshalb neue Konzepte der Evaluation notwendig machten. Prof. Stockmann ging in seinem Vortrag zuerst der Frage nach, was eine ‚gute‘ Wirkungsevaluation ausmacht. Er nannte vor allem folgende Punkte: einen Mehrebenenansatz, Ursache-Wirkungshypothesen, adäquate Wirkungsindikatoren, eine kontinuierliche Datenerfassung, ein dem Wirkungskonzept entsprechendes Untersuchungsdesign, Multimethodenansatz und eine partizipative Vorgehensweise. Danach beschäftigte er sich mit der Frage, warum es erst so wenige ‚gute‘ Impact-Evaluationen gibt. Zum Abschluss nannte der Referent eine Reihe von Gründen, die zu einem vorsichtigen Optimismus Anlaß gäben:

- Nachfrage nach Wirkungsevaluation steige stetig, auch außerhalb der EZ.
- Neue Steuerungsmodelle wie beispielsweise das New Public Management machen Wirkungsevaluationen notwendig.
- Internationaler Druck, Wirkungen von Projekten und Programmen zu belegen, nähme zu.
- Zunehmendes öffentliches Interesse an einer adäquaten Verwendung von Steuermitteln steige.
- Evaluationskompetenzen nehmen weltweit zu. Ein Indikator hierfür seien die vielen Neugründungen von Evaluationsgesellschaften.

Eine gute Wirkungsevaluation benötige hoch spezialisiertes Fachwissen, sei mit einem enormen personellen und zeitlichen Aufwand verbunden und verursache dem-

---

1 <http://www.uni-saarland.de/fak5/stockmann/akepol/>

entsprechend hohe Kosten, vor allem wenn Best Practise Beispiele nicht systematisch zusammengetragen würden und deshalb das Rad immer neu erfunden werde.

Damit solche Reibungsverluste für die Zukunft reduziert werden und eine hohe Glaubwürdigkeit in der öffentlichen Diskussion erreicht wird, sei eine Bündelung der knappen Ressourcen erforderlich. Um dies sicherzustellen empfahl Prof. Stockmann das Aufgreifen einer Idee, die er bereits vor 10 Jahren geäußert hatte, nämlich die Gründung eines nationalen Evaluationsinstituts. Während diese in der Entwicklungszusammenarbeit noch immer belächelt werde, sei sie in anderen Politikfeldern bereits umgesetzt worden.

Mittlerweile wurden u.a. ein Institut für Schulqualität sowie ein Institut für Forschungsqualität und – qualitätssicherung gegründet. Daneben ist auf internationaler Ebene die Idee eines ‚Impact Evaluation Clubs‘ lanciert worden, die der alten ‚Stockmannschen Idee‘ sehr nahe kommt.

Als nächster Referent stellte Herr Peter Maats das methodische Konzept für die Durchführung von Länderprogrammevaluationen sowie bisherige Erfahrungen mit diesem Instrument zur Wirkungsbewertung vor. Hierfür wählte er folgende Organisationen aus: Weltbank (WB), Europäische Kommission (EC) und Department for International Development (DFID). Die präsentierten Ergebnisse wurden im Rahmen einer Studie für das BMZ gewonnen. Gemeinsam sei allen drei Institutionen, dass sie erhebliche Anstrengungen auf institutioneller und methodischer Ebene unternommen haben, den neuen Herausforderungen, die sich mit der Bündelung der Projektarbeit zu Länderprogrammen (LPs) und deren Evaluation ergeben haben, zu begegnen. Gemeinsam ist ihrer Arbeit aber auch, dass sie bisher u.a. aufgrund eines nicht adäquaten Monitoringsystems und schlecht formulierten Wirkungsindikatoren keine Aussage über ihren Beitrag zur Erreichung der MDGs liefern könnten. Die Implementierung eines angepassten Monitoringsystems sowie die Formulierung adäquater Indikatoren werden aller Voraussicht nach die ‚Hauptbaustellen‘ der nächsten Jahre sein, damit LPEs wirkungsorientiert durchgeführt werden können.

Im dritten Vortrag referierte Herr Dr. Pedro Morazán stellvertretend für Frau Dr. Maria Tekülve vom Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) zu ‚Themen, Fragen und Methoden von BMZ-Länderprogrammevaluationen – ein Bericht aus der Werkstatt‘. In einem zweiten Teil seines Vortrags gab Herr Dr. Pedro Morazán einen Überblick über den Stand der Länderprogrammevaluation Honduras, die von ihm als freier Gutachter durchgeführt wurde.

Das BMZ hat erst kürzlich mit der Durchführung von Länderprogrammevaluationen in Honduras, Bangladesh und Uganda begonnen, die zum Zeitpunkt der Tagung noch nicht abgeschlossen waren. Herr Dr. Pedro Morazán führte aus, dass sich in den letzten Jahren sowohl durch die veränderte Schwerpunktsetzung der deutschen EZ in den Partnerländern (Sekor- und Länderprogramme) als auch einer angestrebten besseren Verzahnung der deutschen (‚EZ aus einem Guss‘) sowie der bi- und multilateralen Hilfe, die Anforderungen an Evaluationen verändert hätten. Diese Veränderungen beträfen vor allem die Untersuchungsobjekte, die Themen, die Perspektive, die Methodik sowie die Adressaten von Evaluationen. Im Anschluss daran erläuterte Herr Dr. Pedro Morazán, wie sich das neue Evaluationskonzept im Feld in Honduras bewährt hatte. Im Vordergrund stand nicht die Wir-

kungsbetrachtung des Länderkonzeptes für Honduras, sondern die Erprobung des methodischen Ansatzes selbst.

Den zweiten Tag der Tagung begann Herr Dieter Zürcher mit der Präsentation einer sehr umfangreichen und detaillierten Querschnittsanalyse von Landesprogramm-evaluationen verschiedener Geber, wie der schweizerischen Direktion für Entwick-lung und Zusammenarbeit, der Danish International Development Agency (Dani-da), Swedish International Development Cooperation Agency (Sida) und des De-partment for International Development (DFID). Mit Hilfe dieser Querschnittsana-lyse leitete Herr Zürcher umsetzungsorientierte Schlussfolgerungen für die zukünf-tige Evaluation von Kooperationsstrategien der DEZA, insbesondere bezüglich fol-gender Themenpunkte ab:

- Messung, Definition, Aggregation von Resultaten in Korrelation mit den PRSPs/MDGs korreliert.
- Methoden zur Analyse und Bewertung von Wirkungsketten auf der Ebene Sektor/Land.
- Aufwand und Ertrag dieser Evaluationen.
- Herausarbeitung der Stärken und Schwächen des DEZA Ansatzes im Querver-gleich mit anderen Gebern.
- Herausstellung der wichtigsten Folgerungen für die Planung von neuen Koope-rationsstrategien.
- Minimale Standards für LPEs.

In Bezug auf die Messung von Outputs und Wirkungen auf der Sektor- und Lan-desprogrammebene kam Herr Dieter Zürcher zu ähnlichen Ergebnissen wie Herr Peter Maats. Auch den von ihm untersuchten Institutionen fehlt es noch an adäquan-ten Indikatoren und Daten sowie komplexen Wirkungsketten für eine angemessene Wirkungsbewertung.

Einen Platz für das ‚Narrative in Programmevaluationen‘ forderte Herr Pierre Walther in seinem Vortrag zum Thema ‚Impact Stories‘. Zur Unterstützung seiner Forderung führte Herr Pierre Walter folgende drei Hauptargumente ins Feld:

1. Methodisches Argument: Wirkung von Geschichten.
2. Logisches Argument: Beobachtungen der direkt Betroffenen sind eine Quelle im Diagnoseprozess (Abduktion).
3. Ontologisches Argument: Betroffene Akteure haben Zugang zur Kausalität zwischen Intervention und Wirkung.

Herr Pierre Walter betonte, dass Geschichten anders wirken als Fakten und die Be-reitschaft von Akteuren erhöhen würden, Änderungen im eigenen Verhalten zuzu-lassen. Des Weiteren seien sie kulturell verständlich und hätten eine verbindende Wirkung zwischen z.B. Felderfahrung und Policy Dialog. Geschichten böten dar-über hinaus die Möglichkeit, etwas ‚durch die Blume zu sagen‘, Prozesse zu be-schleunigen und besäßen eine starke mobilisierende Wirkung. Zur Dokumentation und Verarbeitung im Evaluationskontext auf Programmebene präsentierte der Refe-rent eine spezielle Story-Spotting Software.

In ihrem anschließenden Referat stellte Frau Claudia Conrad das Instrument der Real-Time Evaluierungen im Kontext der humanitären Hilfe vor. Real-Time Eva-luierungen (RTE) werden vom UNHCR (2002) wie folgt definiert: „A real-time

Evaluation is a timely, rapid and interactive peer review of fast evolving humanitarian operations undertaken at an early stage. Its broad objective is to gauge the effectiveness and impact of a given response, and to ensure that its findings are used as an immediate catalysts of organizational and operational change.“

RTE würden in einem laufenden Programm zu einem frühen Zeitpunkt durchgeführt, wobei die Ergebnisse der RTE direkt in den Programmprozess einfließen sollten. RTEs zeichneten sich durch einen großen Einbezug und Austausch zwischen Akteuren im Headquarter und im Feld aus und dauern nur wenige Tage. Um eine unverzügliche Einbettung der Evaluationsergebnisse in den Programmprozess zu gewährleisten, müsse die Berichterstattung sozusagen in ‚Real-Time‘, nämlich innerhalb von 4-5 Tagen nach Missionsende, vorliegen. Somit läge der Schwerpunkt der RTEs auch eher auf ‚Lessons Learnt‘ als auf ‚Accountability‘. Evaluatoreninnen und Evaluatoren fungierten im Rahmen von RTEs eher als unterstützende Kraft im Programm (Erleichterung in Stresssituationen) selbst, als dass sie extern und unabhängig agieren würden. Neben dem UNHCR seien wichtige Anwender von RTEs das World Food Programme, UNICEF, Humanitarian Accountability, CARE, World Vision, OXFAM, IFRC und FAO. Die entscheidenden Unterschiede zu einer Programmevaluation im klassischen Sinne sind die, dass die Hauptunterstützungsleistungen von RTEs schon in der Implementierungsphase laufen und folglich keine ex-post Empfehlungen abgegeben würden, wodurch der Fokus natürlich auch auf dem Programmprozess selbst liege und nicht auf der Wirkungsebene anzusiedeln sei. Die praktische Durchführung von RTEs stellte Frau Claudia Conrad beispielhaft an einem Einsatz in Darfur vor.

Den Abschluss der Tagung bildete Frau Monica Rubiolo vom schweizerischen Staatssekretariat für Wirtschaft (seco) mit einem Bericht über die ersten Erfahrungen der seco mit Evaluationen von Budgethilfe. Budgethilfe, so Frau Monica Rubiolo, bedeute einen direkten Finanztransfer in das Staatsbudget von ‚reformorientierten Ländern‘, die über die Verwendung von Geldern, unter Einhaltung bestimmter ausgehandelter Vorgaben (Performance Assessment Frameworks, PAFs), eigenständig bestimmen könnten. Die finanziellen Transferzahlungen der seco seien in den Politikdialog um Strukturreformen und die nationalen Entwicklungsprioritäten eingebettet. Bei der Budgethilfe handele es sich um ein Instrument der EZ, welches in Abstimmung mit den verschiedenen Gebern der jeweiligen Partnerländer durchgeführt wird. Die methodische Grundlage der Evaluationen von Budgethilfe bildet für die seco das ‚Enhanced Evaluation Framework‘. Abschließend zog Frau Monica Rubiolo ein Fazit über die ersten Erfahrungen mit der Evaluation von Budgethilfeprogrammen:

- Länderanalysen hätten sehr viel Potential, die in der Synthese/Empfehlungen/Schlussfolgerungen ungenutzt bleiben.
- Diese Art von Evaluationen zeigten Grenzen einer wissenschaftlich fundierten Impact-Forschung auf: einfache Antworten auf die Frage, ob die Armut reduziert wird, sind nicht erhältlich.
- Die Evaluation liefere jedoch viele interessante Einsichten und Erkenntnisse, sowie relevante konkrete Hinweise, um diese Frage zu beantworten. Sie müssen aber deutlicher und kohärenter dargestellt werden.

Die gemeinsame Frühjahrstagung wurde von den Teilnehmerinnen und Teilnehmer als Erfolg gewertet. Die deutschen und die schweizerischen Evaluationsexpertinnen und -experten in der Entwicklungszusammenarbeit haben beschlossen, den gemeinsamen Erfahrungsaustausch fortzusetzen.