Wirkungsmessung in der Entwicklungszusammenarbeit

Rigorous Impact Evaluation

Evaluation hat in der Entwicklungszusammenarbeit bereits eine lange Tradition. In der Regel werden dabei allerdings Methoden angewandt, die nicht den gängigen Standards in wissenschaftlichen Untersuchungen entsprechen. Ein neuer Ansatz möchte zu einer besseren und effektiveren Evaluationspraxis beitragen.

Von Wolfgang Meyer, Nicolà Reade und Reinhard Stockmann

In der Entwicklungspolitik haben die Millenniums-Erklärung und die Millenniumsentwicklungsziele (MDGs) der Vereinten Nationen sowie die Erklärung von Paris das internationale Evaluierungsgeschehen in den letzten Jahren geprägt. Damit einher geht eine zunehmende Orientierung auf Wirkungen, welche Evaluationen die Aufgabe zuweist, zuverlässige Ergebnisse über alle mit Entwicklungsmaßnahmen bewirkten Veränderungen zu liefern und sich nicht nur auf die Überprüfung der Zielerreichung zu beschränken.

Defizite der Evaluationspraxis

Allerdings entspricht die in der bisherigen Evaluierung übliche Vorgehensweise nicht den hohen methodischen Anforderungen an eine Wirkungsmessung. Anspruchsvolle Erhebungsdesigns wie beispielsweise Experimentaldesigns, Längsschnittstudien oder Multimethodenansätze werden bisher selten verwendet (1). Die gegenwärtige internationale Debatte über methodische Defizite der Evaluationspraxis in der Entwicklungszusammenarbeit wurde durch eine Publikation des Centers for Global Development angestoßen (CDG 2006). Sie wird nun unter dem Stichwort „(Rigorous) Impact Evaluation" geführt (Caspari/Barbu 2008).

Wirkungsevaluationen sehen sich vor zwei Herausforderungen gestellt. Zum einen sollen sie Aufschluss geben über möglichst alle Veränderungen, die sich während der Durchführung von Maßnahmen im Operationsgebiet ergeben haben. Zum anderen sollen diesen festgestellten Entwicklungen kausal Ursachen zugeschrieben werden, wobei natürlich den durchgeführten Interventionen besonderes Augenmerk gilt. Wirkungen umfassen somit die intendierten, nicht-intendierten, positiven wie negativen, erwarteten oder unerwarteten Veränderungen. Zudem kann zwischen internen, bei der Durchführungsorganisation ausgelösten Effekten, und externen, in den Politikfeldern der Intervention und bei den Zielgruppen hervorgerufenen Wirkungen, unterschieden werden.

Ein theoriebasierter Ansatz zur Wirkungsmessung ist somit die Voraussetzung für eine Ordnung der Einflussfaktoren und Wirkungsfelder. Hierzu muss ein Modell über die erwarteten Ursache-Wirkungs-Zusammenhänge aufgestellt und dann anhand der gewonnenen empirischen Daten überprüft werden. Ein Beispiel für ein solches Modell ist in Stockmann (2006b) dargestellt. Evaluationen, die zwar Wirkungen nachweisen, aber keine Angaben machen, warum eine Maßnahme die erwarteten Wirkungen gezeigt hat oder nicht, sind nur von begrenztem Nutzen (White 2006; Ravallion 2005). Die Frage der Kausalität sollte im Zuge der Lernprozesse für die effektive Gestaltung von Entwicklungsmaßnahmen im Vordergrund stehen (Bamberger et al. 2006; Baker 2000).

Ansatz zur Wirkungsmessung

Für die Nachhaltigkeitsdiskussion sind speziell die Differenzierungen zwischen kurz-, mittel- oder langfristigen Wirkungen, zwischen ökonomischen, ökologischen und sozialen Systemfeldern sowie zwischen der gesellschaftlichen Makro-, der gruppen- und organisationenbezogenen Meso- und der auf Individuen orientierten Mikroebene von Bedeutung (Meyer 2007). Methodisch gesehen wird hier der Einsatz von systembezogenen, dynamischen Mehrebenenmodellen erforderlich.

Die gegenwärtige Diskussion der Durchführungsorganisationen orientiert sich allerdings an einfacher formulierten Wirkungsketten. Der Begriff Impact umfasst hier ausschließlich langfristige, übergeordnete entwicklungspolitische Wirkungen wie zum Beispiel die Reduktion der Armut. Darüber hinaus wird zwischen Input, den Interventionsmaßnahmen eines Projektes, dem Output, den Ergebnissen dieser Maßnahmen, und dem Outcome, den kurz- und mittelfristigen, unmittelbaren Wirkungen der Ergebnisse, unterschieden. Im Unterschied zum Input-Output-Vergleich sollen Wirkungsevaluationen nicht nur auf der Mikroebene Effekte aufzeigen und Aufschluss über die Zielerreichung einer Interventionsmaßnahme geben, sondern schwerpunktmäßig ihre Wirkungen auf der Makroebene nachweisen und so beispielsweise den Beitrag eines Projektes oder Programms zur Armutsreduzierung innerhalb einer Region belegen (OECD/DAC 2002).

Die besondere Herausforderung von Wirkungsevaluationen ist somit die Identifikation von Nettowirkungen auf der Makroebene. Es muss dabei insbesondere der Frage nachgegangen werden, was ohne das Projekt oder Programm geschehen wäre. ➜

„Die Diskussion zu Rigorous Impact Evaluation fordert eine stärkere Wissenschaftlichkeit in einem vorwiegend pragmatisch orientierten Politikfeld“

Die Wirkungsmessung muss also das Kontrafaktische berücksichtigen. Hierzu müssten Datenerhebungsdesigns verwendet werden, die Ursache-Wirkungs-Zusammenhänge isolieren und sämtliche potenziellen Störfaktoren kontrollieren können. Eine solche experimentelle Vorgehensweise ist jedoch aufgrund der spezifischen Rahmenbedingungen aus ethischen, politischen und ökonomischen, angesichts der zu erwartenden Komplexität der Wirkungsmodelle aber auch aus methodischen Gründen kaum realisierbar.

Wirkungsevaluationen auf Makroebene

Dementsprechend sind unterschiedliche Designalternativen und deren Umsetzungsmöglichkeiten in der Entwicklungspolitik ein zentrales Thema methodischer Diskussionen. Ziel ist die Konstruktion von Kontroll- oder Vergleichsgruppen, die Rückschlüsse auf das Kontrafaktische zulassen. Das Spektrum der diskutierten Methoden reicht von ökonometrischen Schätzverfahren, wie zum Beispiel bei Augsburg (2006), bis zu qualitativen und stark partizipativ orientierten Vorgehensweisen (Neubert 2000). Dies bildet zugleich die unterschiedlichen Ansprüche von einer primär wissenschaftlichen bis zu einer vorrangig pragmatischen Orientierung bei der Designkonstruktion ab.

Die Diskussion zu Rigorous Impact Evaluation fordert eine stärkere Wissenschaftlichkeit in einem bisher vorwiegend pragmatisch orientierten Politikfeld. Dieser Herausforderung stellen sich zunehmend die großen staatlichen Durchführungsorganisationen in der Entwicklungszusammenarbeit. In Deutschland hat sich zum Beispiel die Gesellschaft für technische Zusammenarbeit vor vier Jahren entschlossen, ihr Evaluationssystem grundlegend zu reformieren und jährlich 30 Evaluationen in zwei verschiedenen Tätigkeitsfeldern von unabhängigen Instituten nach einem neuen, einheitlichen Kriterienkatalog durchführen zu lassen. Diese Fremdevaluationen der GTZ dienen nicht ausschließlich der allgemeinen Erfolgsbewertung, sondern sollen Wirkungsmessungen im Fokus haben. Bisher konnten jedoch die Ansprüche einer Rigorous Impact Evaluation in diesem Rahmen nicht verwirklicht werden, weshalb die GTZ zurzeit an einer Weiterentwicklung ihres Systems arbeitet.

Anmerkungen

(1) Vgl. für Deutschland beispielsweise die Befunde der Systemprüfungen 1999, 2001 und 2007, Borrmann 1999, 2001; als Überblick Stockmann 2006a. Die Publikation der Ergebnisse zur Systemprüfung 2007, die von HWWA und CEval gemeinsam durchgeführt wurde, ist für Anfang 2009 vorgesehen.

Literatur

Augsburg, B.: Econometric Evaluation of the SEWA Bank in India. Applying Matching Techniques based on the Propensity Score, Maastricht 2006.

Baker, J. L.: Evaluating the Impact of Development Projects on Poverty: A Handbook for Practitioners. Washington 2000.

Bamberger, M./Rugh, J./Mabry, L.: Real World Evaluation. Working under Budget, Time, Data And Policy Constraints. Newbury Park 2006.

Borrmann, A. et al.: Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Baden Baden 1999.

Borrmann, A. et al.: Reform der Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Baden Baden 2001.

Caspari, A. / Barbu, R.: Wirkungsevaluierungen: Zum Stand der internationalen Diskussion und dessen Relevanz für Evaluierungen der deutschen Entwicklungszusammenarbeit. Bonn 2008.

CDG: When Will We Ever Learn? Improving Lives through Impact Evaluation. Washington, D.C 2006.

Meyer, W.: Vertical Dimension of Social Integrations. Grass-root Activities for Managing Sustainability. In: Rati, R.N.R / Schwarz-Herion, O. (Eds.): Sustainable Development: Issues and Perspectives. New Delhi 2007, S. 99 - 123.

Neubert, S.: Social Impact Analysis of Poverty Alleviation Programmes and Projects. A Contribution to the Debate on the Methodology of Evaluation in Development Cooperation, London 2000.

OECD/DAC: Glossary of Key Terms in Evaluation and Results Based Management. Paris 2002

Ravallion, M.: Evaluating Anti-Poverty Programs. Washington, D.C. 2005.

Stockmann, R.: Evaluation staatlicher Entwicklungspolitik. In: Stockmann, R. (Hrsg.): Evaluationsforschung. Münster 2006a.

Stockmann, R.: Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Münster 2006b.

White, H.: Impact Evaluation: An Overview And Some Issues For Discussion. 2006.

AUTOREN + KONTAKT

Dr. Wolfgang Meyer ist stellvertretender Leiter und Koordinator des Bereichs „Arbeit und Umwelt“ am CEval,
Nicolà Reade arbeitet zu Entwicklungszusammenarbeit und Wirkungsevaluation am CEval und
Dr. Reinhard Stockmann ist Leiter des Instituts CEval und Lehrstuhlinhaber an der Universität des Saarlandes.

CEval - Centrum für Evaluation,
Universität des Saarlandes, Postfach 15 11 50,
66041 Saarbrücken. Tel.: +49 681 302,
E-Mail: info@ceval.de

(c) 2010 Authors; licensee IÖW and oekom verlag. This is an article distributed under the terms of the Creative Commons Attribution Non-Commercial No Derivates License (http://creativecommons.org/licenses/by-nc-nd/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.