

## Editorial

Staatliches Handeln, das durch Bundes- und Landesministerien, Kommunen oder öffentliche Einrichtungen gesteuert wird, genauso wie unzählige zivilgesellschaftliche Aktivitäten, die über eine Vielzahl von Nicht-Regierungsorganisationen entfaltet werden, sind auf die Veränderung bestehender Verhältnisse ausgerichtet. Einzelne Maßnahmen, Projekte, Programme und Entwicklungs-, Reform- oder Innovationsstrategien dienen dazu, als wünschenswert erklärte Ziele zu erreichen, womit nichts anderes als geplante Veränderungen gemeint sind: die intendierten Wirkungen dieser Interventionen. Ob sie tatsächlich eintreten, bleibt oft im Dunkeln. Stattdessen begnügen sich viele staatliche wie nicht-staatliche Akteure noch immer gerne damit, auf die investierten personellen und materiellen Ressourcen zu verweisen (Inputs) und die erbrachten Leistungen (Outputs) zu messen.

So erfährt die Öffentlichkeit schnell etwas über die Zahl neu eingestellter Lehrer(innen) oder Polizeibeamter bzw. -beamtinnen (Input), aber nichts darüber, ob sich dadurch die Unterrichtsqualität oder die Aufklärungsquote von Verbrechen erhöht hat. Auch die Zahl geleisteter Unterrichtsstunden oder Streifenfahrten (Output) sagt nichts über die eigentliche Zielerreichung aus – vorausgesetzt, dass diese sich nicht in der Erhöhung des Inputs oder Outputs erschöpft, sondern der Verwirklichung eines erklärten Zieles dient, z.B. der Steigerung des Wissensniveaus von Schüler(inne)n oder der Aufklärungsquote von Verbrechen.

In der Entwicklungszusammenarbeit (EZ) war es bisher nicht anders als in den meisten Politikfeldern. Doch wie kaum in einem anderen Bereich ist die Wirkungsorientierung in den letzten zehn Jahren zu einem dominanten Handlungsprinzip in der EZ avanciert. Hierzu haben auf internationaler Ebene zum einen die Millenniums-Deklaration und die Millenniumsziele (2000) einen wichtigen Beitrag geleistet, die mit einem konkreten Zielrahmen bis 2015 sowie einem umfangreichen Monitoring- und Evaluationssystem zur Überprüfung der Zielerreichung ausgestattet sind. Zum anderen hat eine international geführte Aid Effectiveness Debatte, verstärkt durch die sogenannte Pariser Erklärung (2005) über die Wirksamkeit der EZ sowie die Accra Agenda for Action (2008), den Druck, sich mit den Wirkungen

von Interventionen auf den verschiedensten gesellschaftlichen Ebenen zu beschäftigen, erhöht.

Auch in Deutschland hat die Debatte um eine mangelhafte Wirkungsmessung in der EZ erste Früchte getragen. Bereits 2002 wurde zwischen dem Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) und der Gesellschaft für Technische Zusammenarbeit (GTZ) ein neuer Auftragsrahmen (AURA) vereinbart, der die entwicklungspolitische Wirkung zum Handlungsprinzip bei der Planung und Durchführung von Programmen sowie zum Hauptkriterium bei Evaluationen erhoben hat. Mittlerweile haben die meisten staatlichen Durchführungsorganisationen diesen neuen Auftragsrahmen übernommen und auch viele nicht-staatliche EZ-Organisationen haben die Wirkungsorientierung in den Mittelpunkt ihrer Arbeit gestellt.

Welche Folgen dies für die Evaluation von EZ-Projekten und -Programmen hat, mit welchen Evaluationsdesigns und Methoden die anspruchsvolle Aufgabenstellung intendierte wie nicht-intendierte Wirkungen zu messen und ihren Ursachen zuzuschreiben, möglichst optimal gelöst werden kann, ist Gegenstand vielfältiger Diskussionen, Workshops und Tagungen. Zudem wird in einer Reihe von Studien bereits damit experimentiert, wie die Evaluation der Zukunft in der EZ aussehen könnte.

Darum geht es schwerpunktmäßig in diesem Heft. Ausgehend von einem Workshop, den das BMZ gemeinsam mit dem Deutschen Institut für Entwicklungspolitik (DIE), im September 2008 in Bonn, zu dem Thema „Rigorese Wirkungsanalyse in der EZ“ veranstaltete, entstand die Idee, eine kleine Zahl von Studien einer breiteren Fachöffentlichkeit vorzustellen. Ein herzlicher Dank gilt in diesem Zusammenhang Ministerialrätin *Michaela Zintl*, Leiterin des BMZ-Referats Evaluierung der Entwicklungszusammenarbeit/Außenrevision, die diesen Workshop initiiert, organisiert und geleitet hat, und ohne deren Engagement dieses Heft nicht zustande gekommen wäre.

Im ersten Beitrag dieses Hefts beschäftigt sich *Alexandra Caspari* mit der Debatte um „rigorese“ Wirkungsevaluationen. Damit sind u.a. Evaluationsdesigns gemeint, die nicht nur einen Vorher-Nachher-Vergleich erlauben, sondern bei denen neben einer Versuchsgruppe (Zielgruppe, bei der eine Intervention stattfindet) auch eine Kontrollgruppe zum Vergleich herangezogen wird. Solche experimentellen oder quasi-experimentellen Designs bieten die beste Gewähr dafür, relevante Störgrößen, also andere Faktoren, die für die Erklärung festgestellter Wirkungen verantwortlich gemacht werden könnten, auszuschalten. Die Autorin hält rigorese Wirkungsevaluationen vor allem dann für sinnvoll, wenn die untersuchten Programme ein hohes Lernpotential aufweisen. Dies sieht Caspari (2009: 210), insbesondere dann gegeben, „wenn eine ‚Policy‘ oder ein Programm von *strategischer Relevanz* für die Armutsreduzierung ist, wenn ein *neuer, vielversprechender, innovativer Ansatz* oder auch eine *neue Policy* eingeführt wurde oder wenn ausreichend Kenntnisse vorhanden sind, dass die untersuchte Maßnahme auch in *anderen Kontexten* wirksam ist“.<sup>1</sup>

---

1 Caspari, Alexandra (2009): ‚Rigorese‘ Wirkungsevaluation – methodische und konzeptionelle Ansätze der Wirkungsmessung in der Entwicklungszusammenarbeit. In: Zeitschrift für Evaluation, Jg. 8, Bd. 2/2009, S. 183-213.

*Jan R. Böhnke, Jan Koehler und Christoph Zürcher* erproben in ihrer, im Auftrag des BMZ durchgeführten Evaluation „rigorose“ Methoden, um den Einfluss der EZ auf die Friedenssicherung in Post-Konflikt-Zonen (Afghanistan) zu messen und kausal zu erklären. Sie haben hierfür ein aufwändiges Stichprobenverfahren angewendet, um Ziel- und Vergleichsgruppen voneinander zu unterscheiden. Es sind über die erste Erhebung im Jahr 2007 hinaus, weitere in zweijährigem Abstand geplant und es werden qualitative und quantitative Datenerhebungsverfahren eingesetzt. Dabei kommen die Autoren zu dem zunächst nicht weiter überraschenden Befund, dass EZ-Maßnahmen einen Einfluss auf die Einstellungen gegenüber internationalen Akteuren sowie der staatlichen Legitimität haben, aber auch, dass die EZ-Maßnahmen keinen Einfluss auf die Sicherheitswahrnehmung der Befragten haben!

Über eine andere Herangehensweise berichtet *Nicolà Reade* in ihrem Beitrag über Pilot-Wirkungsevaluationen im Wassersektor. Bei diesem Versuch, die im Auftrag der GTZ durchgeführten, unabhängigen Evaluationen rigoroser, d.h. mit adäquateren Designs und Methoden, zu untersuchen, wird die GTZ-Wirkungskette als Grundlage für eine hypothesengeleitete Ursache-Wirkungs-Analyse verwendet, was nicht unproblematisch ist, da diese Kette nicht (immer) auf systematischen Wenn-Dann-Aussagen beruht. Dennoch kommt die Autorin u.a. zu dem Ergebnis, dass quasi-experimentelle Designs in der EZ einsetzbar und kausale Ursachenzuschreibungen möglich sind, dass repräsentative, standardisierte Ziel- und Vergleichsgruppenbefragungen für die Kausalanalyse essentiell sind und retrospektive Befragungen in bestimmten Fällen Baselines ersetzen können.

Der Beitrag von *Reinhard Stockmann und Axel Borrmann* beschäftigt sich mit dem Evaluationssystem der EZ als Ganzes. Im Auftrag des BMZ wurde unter ihrer Leitung untersucht, inwieweit das BMZ selbst, die deutschen Durchführungsorganisationen und eine Reihe zivilgesellschaftlicher Organisationen ihre Evaluationen auf die eingangs beschriebenen Veränderungen in der nationalen und internationalen Förderlandschaft eingestellt haben. Aus den 20 detaillierten Fallstudien und der sich anschließenden Metaanalyse wird deutlich, dass zwischen den untersuchten Organisationen in konzeptioneller, organisatorischer und methodischer Sicht eine enorme Heterogenität besteht, die sich in den letzten zehn Jahren weiter vergrößert hat. Eine verstärkte Systembildung und weitreichende Qualitätsverbesserungen werden als dringend erforderlich empfohlen.

Im Informations- und Serviceteil dieses Heftes stellt *Sebastian Elbe* den „Evaluierungs-Helpdesk des Europäischen Evaluierungsnetzwerkes für ländliche Entwicklung“ vor. Die von der Generaldirektion (GD) Landwirtschaft unterstützte Einrichtung soll u.a. die Netzbildung fördern, die Methoden- und Instrumentenentwicklung vorantreiben und gute Praxisbeispiele von Evaluationen ländlicher Entwicklungsprogramme sammeln.

In der Reihe „Evaluationsinstitute stellen sich vor“ präsentiert sich dieses Mal die Abteilung „Qualität und Evaluierung“ der Internationalen Weiterbildung und Entwicklung gGmbH (InWEnt). Drei interessante Rezensionen runden den Serviceteil dieses Heftes ab.

Nützliches und Anregendes finden Sie auch in der Rubrik *DeGEval ...Info*. Abwechslungsreiche Informationen zu den vielfältigen Aktivitäten der Arbeitskreise der DeGEval – Gesellschaft für Evaluation sowie ein lesenswerter Beitrag zum Thema „Evaluation und Steuerung“ schließen diesen Band ab.

Allen, die dazu beigetragen haben, diese interessante Ausgabe zu gestalten, danke ich herzlich und wünsche allen Leserinnen und Lesern eine anregende Lektüre.

*Reinhard Stockmann*  
Für die Herausgeberinnen und Herausgeber