DEval
DEUTSCHES EVALUIERUNGSINSTITUT DER ENTWICKLUNGSZUSAMMENARBEIT

DEval Policy Brief 02/2016

„SYSTEMPRÜFUNG“ – EIN IMPULSGEBER FÜR DIE EVALUIERUNGSPRAXIS IN DER DEUTSCHEN ENTWICKLUNGSZUSAMMENARBEIT?

Zusammenfassung

Belastbare Aussagen zur Wirksamkeit der deutschen Entwicklungszusammenarbeit (EZ) sind von zentraler Bedeutung – sei es für Rechenschaftslegung, sei es für verbesserte Projekt- oder Programmsteuerung oder für konzeptionelles Lernen. In Ergänzung zu wirkungsorientierter Planung und zum Monitoring gilt unabhängige Evaluierung im OECD/DAC-Geberkreis als das Mittel der Wahl, um die hierfür notwendigen Erkenntnisse zu gewinnen.

Um den Stand der Evaluierungspraxis in der deutschen EZ in Erfahrung zu bringen, veranlasste das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) Anfang 2007 eine sogenannte Systemprüfung (vgl. Borrmann und Stockmann, 2009). Als Ergebnis wurden neben spezifischen Handlungsempfehlungen für die untersuchten Entwicklungsorganisationen zur Weiterentwicklung ihrer organisationsinternen Evaluierungspraxis auch Empfehlungen auf Systemebene formuliert. Eine davon war die Gründung eines unabhängigen Evaluierungsinstituts. Diese Empfehlung wurde 2012 mit der Gründung des Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) umgesetzt.

Die Auswirkungen der Systemprüfung auf die Evaluierungspraxis in der deutschen EZ sowie darüber hinausgehende Entwicklungen der Evaluierungspraxis hat das DEval im Jahr 2014 in einer Studie untersucht. Dabei wurden folgende Erkenntnisse gewonnen:

- Das Instrument der Systemprüfung hat sich bewährt. Es hat sowohl in den untersuchten Organisationen als auch auf Systemebene wichtige Impulse zur Weiterentwicklung der Evaluierungspraxis gesetzt. Es empfiehlt sich daher, solche Überprüfungen des Evaluierungssystems auch künftig durchzuführen.
- Die Institutionalisierung der Evaluierungspraxis ist erheblich vorangeschritten, und die Aufgeschlossenheit gegenüber Evaluierung hat sich erhöht. Dies spiegelt sich bei den Organisationen unter anderem in erhöhten (Personal-) Ressourcen für Evaluierung sowie in gestärkten Evaluierungsstrukturen und Prozessen (vermehrter Einsatz von Evaluierungsrichtlinien und -plänen, standardisierte Prozesse, bessere institutionelle Stellung der Evaluierungseinheiten) wider.
- Als wichtigste Zielsetzung von Evaluierungen nennen die meisten EZ-Organisationen die Verbesserung der Vorhaben und das (instrumentelle) Lernen. Gewonnene Erkenntnisse verbleiben jedoch häufig bei den direkt am Projekt Beteiligten, und Evaluierungsberichte werden immer noch selten vollständig veröffentlicht. Um das Ausschöpfen von konzeptionellem und interinstitutionellem Lernpotenzial zu fördern, sollte verstärkt eine Kultur des „transparenten Lernens“ geschaffen werden.
- Eine zentrale Herausforderung für viele Evaluierungen besteht nach wie vor im Einsatz methodisch anspruchsvoller Designs, die auf nachvollziehbaren Wirkungshypothesen basieren. Diesbezüglich besteht seitens der EZ-Organisationen auch ein starkes Interesse an Qualifizierungsmaßnahmen.
- Die systematische Einbindung der Partner in den Evaluierungsprozess ist häufig ausbaufähig. Hier bestehen erhebliche Unterschiede zwischen den verschiedenen EZ-Organisationen.
- Die vom BMZ verantwortete EZ wird umfassend evaluiert. Diese Evaluierungsdichte fehlt anderen Ressorts, die öffentliche Entwicklungsleistungen umsetzen, aber auch anderen Politikbereichen.

Die Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit

Ob Millenniums-Entwicklungsziele, Paris-Deklaration oder Agenda 2030 für nachhaltige Entwicklung – all diese ambitionierten globalen Zielsetzungen verdeutlichen die Notwendigkeit einer wirkungsvollen, nachhaltigen und kohärenten Entwicklungspolitik. Evaluierungen liefern hierzu evidenzbasierte Erkenntnisse, die von entscheidender Bedeutung für den zukünftigen, zielgerichteten Einsatz von entwicklungspolitischen Maßnahmen sind und der Rechenschaftslegung über den Einsatz der eingesetzten Steuermittel dienen.

Vor diesem Hintergrund wird Evaluierung immer bedeutender – umso mehr, als für die Agenda 2030 ein effektiver gemeinsamer Rahmen für Monitoring und Evaluierung (M&E) gefunden werden muss. Wie steht es nun aber um die Evaluierungspraxis in der deutschen EZ sowie um die Qualität der dabei gewonnenen Erkenntnisse? Dies wurde im Rahmen einer Systemprüfung anhand von Fallstudien staatlicher und nicht-staatlicher EZ-Organisationen untersucht, die das BMZ Anfang 2007 in Auftrag gab (vgl. Box). Dabei wurde deutlich, dass sich die Heterogenität der deutschen EZ-Organisationen auch in deren Evaluierungssystemen widerspiegelt. Eine zentrale Empfehlung auf Systemebene war daher die Gründung eines unabhängigen Evaluierungsinstituts. Dieses sollte das BMZ bei strategischen Entscheidungen durch unabhängige Evaluierungen zur Wirksamkeit, Wirtschaftlichkeit und Nachhaltigkeit der EZ unterstützen (Borrmann und Stockmann, 2009: 17, 65, 165) und so konzeptionelles und interinstitutionelles Lernen auf Systemebene fördern. Mit der Gründung des DEval setzte das BMZ diese Empfehlung um.

Daneben gab die Systemprüfung Aufschluss darüber, wie die verschiedenen EZ-Organisationen ihre mit BMZ-Mitteln finanzierten Maßnahmen evaluieren. Dabei zeigte sich, dass die Aussagekraft dieser Evaluierungen von zahlreichen Aspekten abhängt. So unterscheiden sich Zielsetzung (z.B. Lernen oder Rechenschaftslegung) und Form der durchgeführten Evaluierungen (z.B. Zwischen- oder Expost-, Projekt- oder Querschnittsevaluierung) sowie die verwendeten Designs und Methoden. Diese Merkmale bestimmen wesentlich, wie evidenzbasiert und nützlich die Ergebnisse sind. Darüber hinaus werden die Evaluierungsergebnisse und deren Nutzung durch die Ressourcenausstattung, aber auch durch Strukturen und Prozesse, wie etwa die Unabhängigkeit der Evaluierungsverantwortlichen, Formate für institutionelles Lernen aus Evaluierung oder die Einbindung der Partner beeinflusst.

Das DEval untersuchte, wie sich der aktuelle Stand der Evaluierungspraxis in der deutschen EZ darstellt und welche Entwicklungen sich seit der letzten Systemprüfung ergeben haben. Neben einer Online-Befragung von 34 EZ-Organisationen wurden hierzu Interviews mit den beteiligten EZ-Organisationen der vorigen Systemprüfung geführt. Zudem fand eine Meta-Evaluierung von 45 Evaluierungsberichten statt.

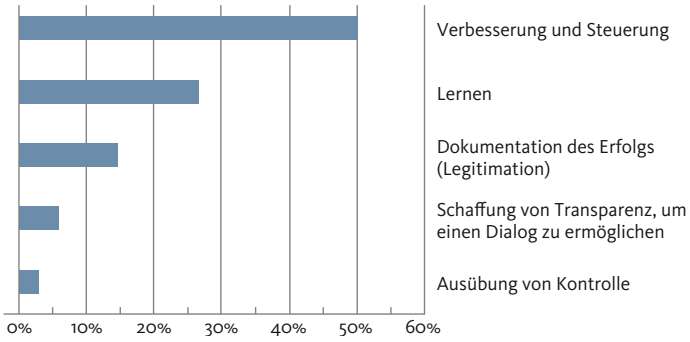
„Systemprüfungen“ in der deutschen EZ

In den Jahren 1997/1998 veranlasste das BMZ erstmals eine systematische Untersuchung der Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit (Borrmann et al., 1999). Eine weitere Untersuchung wurde in den Jahren 2007/2008 durchgeführt (Borrmann und Stockmann, 2009). Hierbei wurden, ausgehend von 20 Einzelfallstudien, systemweite sowie organisationsspezifische Handlungsempfehlungen zur Weiterentwicklung der Evaluierungspraxis ausgesprochen. Im Anschluss wurden zwischen dem BMZ und den beteiligten Organisationen Vereinbarungen zur Umsetzung der von den Organisationen akzeptierten Handlungsempfehlungen getroffen.

Aktueller Stand der Evaluierungspraxis und Impulse durch die Systemprüfung

Als wichtigste Zielsetzung von Evaluierungen gelten für die Mehrheit der befragten EZ-Organisationen die Verbesserung und Steuerung von Vorhaben sowie das Lernen. Die Legitimierung der durchgeführten Maßnahmen, die Schaffung von Transparenz sowie die Ausübung von Kontrolle werden hingegen seltener genannt (vgl. Abbildung 1 sowie Lücking, Freund und Bettighofer, 2015; vgl. auch Borrmann und Stockmann, 2009: 130).

Abbildung 1: Ziele von Evaluierungen



Quelle: Online-Befragung DEval 2014; Einfachauswahl, n=34

Die Organisationen unterliegen auch Evaluierungspflichten gegenüber dem BMZ. In der Praxis werden Ergebnisse aus obligatorischen Evaluierungen hauptsächlich von den unmittelbaren Projektbeteiligten genutzt und in die Programm- oder Projektsteuerung eingespeist. Um projektübergreifendes, konzeptionelles Lernen zu fördern, werden seit der letzten Systemprüfung vermehrt Querschnittsevaluierungen durchgeführt.

Insbesondere die Kontrollfunktion von Evaluierungen kann zu einer ablehnenden Haltung bei den beteiligten Stakeholdern führen. Im Vergleich zur vorigen Systemprüfung zeigt sich jedoch, dass die Aufgeschlossenheit von Mitarbeitenden gegenüber Evaluierungen zugenommen hat: Evaluierungen werden heute als Teil des professionellen Selbstverständnisses aufgefasst, was sich konkret in einer gestiegenen Nachfrage nach Evaluierungen äußert. Die Akzeptanz spiegelt sich aber auch in strukturellen Veränderungen wider: Seit der Systemprüfung sind die finanziellen und personellen Ressourcen für Evaluierung gestiegen.

Die zunehmende Institutionalisierung von Evaluierung zeigt sich an der verstärkten Einführung von Evaluierungsrichtlinien, die sich zunehmend an etablierten nationalen und internationalen Standards wie denen der Gesellschaft für Evaluation (DeGEval) oder des Entwicklungsausschusses der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD/DAC) orientieren. Auch der Austausch von EZ-Organisationen über Evaluierung hat zugenommen und wird von den Beteiligten einvernehmlich als hilfreich eingeschätzt. Hingegen bleibt der Austausch mit wissenschaftlichen Einrichtungen zu Methoden und Designs ausbaufähig.

Welches Entwicklungspotenzial hat die Evaluierungspraxis in der deutschen EZ?

Neben den beschriebenen positiven Entwicklungen gibt es aber auch weiterhin Bereiche mit Entwicklungspotenzial:

In der Systemprüfung wurde kritisiert, dass projektübergreifende sowie Wirkungs- und Expost-Evaluierungen zu selten durchgeführt würden. Obwohl Projektevaluierungen weiterhin den Großteil der durchgeführten Evaluierungen darstellen, ist insbesondere die Anzahl an Querschnittsevaluierungen gestiegen. Ein hemmender Faktor für die Durchführung von Expost- oder Wirkungsevaluierungen ist die Verfügbarkeit von finanziellen und personellen Ressourcen. Zudem bestehen bei diesen Evaluierungsformen Zweifel am Mehrwert und an der Anwendbarkeit robuster Methoden, insbesondere in Anbetracht der hierfür benötigten Ressourcen.

Mit Blick auf belastbare Aussagen zur Wirksamkeit von Vorhaben haben die im Rahmen der DEval-Studie untersuchten Evaluierungsberichte gezeigt, dass es im Evaluierungsdesign häufig an operationalisierten und damit überprüfbaren Wirkungshypothesen fehlt. Wie bereits in der Systemprüfung kritisiert, wird zudem das mögliche Methodenspektrum nicht genutzt: Überwiegend werden qualitative Dokumentenanalysen und Interviews durchgeführt – rund die Hälfte der untersuchten Berichte stützte sich ausschließlich auf eine dieser Methoden beziehungsweise die Kombination von beiden. Deutlich seltener werden quantitative Erhebungen und Analysen vorgenommen, beispielsweise anhand von standardisierten Fragebögen. Entsprechend ausbaufähig ist der Einsatz von Mixed-Methods-Ansätzen sowie einer nachvollziehbaren Triangulation. (Quasi-)experimentelle Designs sind nach wie vor selten.

Teilweise lassen finanzielle Limitationen anspruchsvolle Designs und Methoden nicht zu. Zudem werden Gutachterinnen und Gutachter oft vor allem aufgrund fach- oder landesspezifischer Kenntnisse ausgewählt. Es fehlt bei den Evaluierungen daher häufig an Wissen über belastbare qualitative und insbesondere quantitative Methoden. Im Zuge der Untersuchung wurde vielfach ein Interesse an Qualifizierung in anspruchsvolleren Methoden geäußert.

Projektübergreifendes, konzeptionelles Lernen wird als wichtiges Ziel von Evaluierungen angesehen (vgl. Abbildung 1, Seite 2), weshalb zu diesem Zweck auch zunehmend Querschnittsevaluierungen durchgeführt werden. Allerdings sind Evaluierungsprozesse häufig nicht systematisch auf institutionelles Lernen ausgerichtet. Lernerfahrungen werden oft nur mit den Stakeholdern der evaluierten Projekte und Programme geteilt – und auch dies oft nicht in ausreichendem Maße, z.B. was das Teilen von Evaluierungsberichten mit den Partnern vor Ort sowie deren Einbeziehung in den Evaluierungsprozess betrifft. Zudem wird interinstitutionelles beziehungsweise gemeinsames Lernen dadurch erschwert, dass Evaluierungsberichte nur selten vollständig publiziert werden. Insbesondere aufgrund von Schutzinteressen gegenüber Partnern oder der Sorge um das Spendenaufkommen werden in der Regel nur Kurzfassungen veröffentlicht. Aus Sicht der einzelnen Organisationen bleibt daher der Austausch zu Evaluierungsergebnissen sowohl nach innen als auch nach außen begrenzt. Dies kann das Ausschöpfen bestehender Lernpotenziale sowie die möglichst breite Nutzung von Evaluierungsergebnissen einschränken.

Fazit

Die Systemprüfung hat das Evaluierungssystem der deutschen EZ gestärkt. Durch ihre Empfehlungen wurden Diskussionen und Austausch über Nutzen und Nutzung von Evaluierung angestoßen.

Neben der Gründung eines unabhängigen Evaluierungsinstituts zur Stärkung der Evaluierungskapazitäten für komplexe und strategische Evaluierungen sowie zur Weiterentwicklung von Methoden und Standards wurde die Evaluierungspraxis in vielen EZ-Organisationen ausgebaut und weiterentwickelt – auch wenn weiterhin in vielen Bereichen Raum für Verbesserungen besteht:

- Um evidenzbasiertes, konzeptionelles Lernen zu fördern, ist die Nutzung von Wirkungs-, projektübergreifenden und Ex-post-Evaluierungen weiterhin (finanziell und methodisch) zu stärken. Dabei ist mehr Gewicht auf das Evaluierungsdesign und die angewandten Methoden zu legen.
- Betrachtet man die Zielkonflikte bei Evaluierungen, erscheint es sinnvoll, eine politisch gewollte Kultur des „transparenten Lernens“ zu schaffen, in der auch kritische Evaluierungsergebnisse stärker als bisher veröffentlicht und diskutiert werden.
- Im Verantwortungsbereich des BMZ wird vergleichsweise umfassend evaluiert. Andere Ressorts mit Verantwortung für öffentliche Entwicklungsleistungen, aber auch andere Politikbereiche, könnten diesem Beispiel folgen.
- Weitere Überprüfungen des Evaluierungssystems erscheinen in angemessenen Zeitabständen sinnvoll, um zwischenzeitlich eingetretene Entwicklungen zu erfassen oder neue anzustoßen.

Literatur:

Borrmann, A., K. Fasbender, M. Holthus, A. von Gleich, B. Reichl und R. Shams (1999), *Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit. Studie im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung,* NOMOS-Verlagsgesellschaft, Baden-Baden.

Borrmann, A. und R. Stockmann (2009), *Evaluation in der deutschen Entwicklungszusammenarbeit,* Band 1 – Systemanalyse, Waxmann, Münster.

Lücking, K., S. Freund und S. Bettighofer (2015), *Evaluierungspraxis in der deutschen Entwicklungszusammenarbeit. Umsetzungsmonitoring der letzten Systemprüfung und Charakterisierung wesentlicher Elemente,* Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Bonn.

Autoren

Dr. Kim Lücking
Senior Evaluatorin, DEval

Simon Bettighofer
Ehemals Evaluator DEval

Simon Freund
Evaluator, DEval

Das Deutsche Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) ist vom Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) mandatiert, Maßnahmen der deutschen Entwicklungszusammenarbeit unabhängig und nachvollziehbar zu analysieren und zu bewerten. Mit seinen Evaluierungen trägt das Institut dazu bei, die Entscheidungsgrundlage für eine wirksame Gestaltung des Politikfeldes zu verbessern und die Transparenz zu den Ergebnissen zu erhöhen.

Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval)

Fritz-Schäffer-Straße 26 · 53113 Bonn
Telefon +49 (0)228 33 69 07-0

info@DEval.org
www.DEval.org