



Neue Ansätze in der Entwicklungszusammenarbeit

# Von der Pflicht zur Kür

Ein Plädoyer für den strategischen Nutzen von Monitoring und Evaluation

Angelika Klein / Lukas Kupfernagel

Wenn in der Entwicklungszusammenarbeit von Monitoring und Evaluierung gesprochen wird, dann geht es nach wie vor oft darum, den Einsatz von Mitteln – häufig Steuergeldern – zu rechtfertigen und dessen Wirksamkeit nachzuweisen. Das ist richtig und wichtig, doch Monitoring und Evaluierung können mehr. Das Ziel muss ein Bewusstseinswandel sein: weg von den „Hütern der Indikatoren“ hin zum „Freund und Helfer“.

---

„Kein Politikfeld wird so intensiv auf seine Wirksamkeit hin abgeklopft wie die Entwicklungspolitik.“<sup>1</sup> Diese Aussage von Staatssekretär Martin Jäger spiegelt den Rechtfertigungsdruck wider, dem Entwicklungs- und internationale Zusammenarbeit ausgesetzt sind. Dieser wiederum ist mehr als verständlich: Schließlich muss dem Steuerzahler vermittelt werden, in welchem Umfang und wie zweckgebunden Steuergelder im Ausland verausgabt werden und welche Wirkung sie erzielen. Zudem ist die Entwicklungszusammenarbeit (EZ) nicht immer unumstritten. Wissenschaftler wie der US-Ökonom William Easterly kritisieren Entwicklungshilfearbeit internationaler Institutionen bereits seit Beginn der 2000er Jahre, und zwar vor allem unter ökonomischen Gesichtspunkten: Empfängerstaaten schlugen kontraproduktive Wege ein, wenn falsche Anreize gesetzt würden, und dies sei in den vergangenen Jahrzehnten öfter zu verzeichnen als Erfolgsgeschichten.<sup>2</sup> Und der Druck wächst. Zunehmend stellen auch andere Akteure die steigende Abhängigkeit zahlreicher Empfängerstaaten von Entwicklungshilfegeldern in den Vordergrund.<sup>3</sup> Vertreter populistischer Strömungen etwa beharren auf dem Standpunkt, dass für EZ-Projekte verausgabte Gelder überwiegend Verschwendung seien und gestrichen werden müssten. Und auch in der deutschen Politik werden Diskussionen über Budgetkürzungen, wie sie unter dem ehemaligen US-Präsidenten Donald Trump stattfanden,<sup>4</sup> insbesondere durch die Bundestagsfraktion der AfD geführt.

Die durch die COVID-19-Pandemie zu erwartende finanziell und wirtschaftlich schwierige Lage

wird die Debatte zusätzlich befeuern und zunehmend aus dem politischen in den gesellschaftlichen Raum hineinragen. Die Öffentlichkeit merkt ohnehin schon auf: Entwicklungszusammenarbeit sei teuer, ineffizient und richte im Zweifel mehr Schaden an, als den eigentlichen Zielgruppen zu helfen. Insbesondere der finanzielle Aspekt wird mit Blick auf die Verwendung von Steuergeldern angeführt, um die EZ grundsätzlich infrage zu stellen. Eine Studie im Auftrag des Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) belegt, dass jeder vierte deutsche Bürger EZ für unwirksam und nur jeder zehnte sie für wirksam hält. Die Kritiker bemängeln auch, dass etwa die Hälfte aller EZ-Gelder durch Korruption verloren gehe.<sup>5</sup>

Gleichzeitig steigt der Druck durch internationale Akteure: Immer mehr Staaten, die den Prinzipien der freiheitlich-liberalen Weltordnung ablehnend gegenüberstehen, locken durch lukrative, nicht konditionierte Angebote traditionelle Partner der deutschen EZ und sorgen für „Geber-Konkurrenz“ in den Empfängerländern. Dass diese Verführungen nicht immer von Nachhaltigkeit und dem Interesse der Adressaten, sondern von eigenen finanziellen und strategischen Vorteilen und Interessen geprägt sind, scheint ihnen nicht im Wege zu stehen.<sup>6</sup> Die Frage, ob und wie lange der von Seiten des „Westens“ propagierte Anspruch einer wertgebundenen EZ in Sachen Attraktivität mithalten kann, wird immer banger gestellt werden müssen.

Kurzum: Experten, Politiker und Institutionen, die Entwicklungszusammenarbeit betreiben,

stehen unter zunehmendem Legitimationsdruck. In Deutschland – und nur hierum soll es im Folgenden gehen – muten die Erwartungen an EZ-Akteure dabei zuweilen utopisch an: Maximal und schnellstmöglich soll die Wirkung sein bei gleichzeitig geringen und vollständig nachvollziehbaren Ausgaben. Für die Technische Zusammenarbeit (TZ) kein befremdlicher Anspruch: Gerade die TZ und Katastrophenhilfe sollten sich hieran messen lassen können. Im Governance-Bereich ist die Lage – nicht nur aufgrund des langfristigen Engagements – allerdings komplexer. Dennoch kann und darf sich die EZ der Notwendigkeit, Wirksamkeit nicht nur anzustreben, sondern auch nach außen hin zu dokumentieren und zu kommunizieren, grundsätzlich nicht entziehen. Wie kann dies überzeugend gelingen und welche Instrumente gibt es, um dem – teilweise durchaus begründeten – Rechtfertigungsdruck nicht nur standzuhalten, sondern ihn für sich selbst nutzbar zu machen?

### **Monitoring und Evaluierung als Kompass?**

Jenseits ethischer Prinzipien und Werte sollten Daten bestmöglich belegen, dass finanzielle und personelle Ressourcen gut angelegt sind und die Sinnhaftigkeit der Projekte garantiert ist. Vor allem zwei Instrumente können dazu dienen, das EZ-Schiff auf Kurs zu halten und durch Untiefen zu manövrieren: Monitoring und Evaluierung (M&E). Bislang sind diese Begriffe hauptsächlich zum Zweck der Rechtfertigung genutzt worden, doch sie bieten mehr. Nämlich die Möglichkeit, proaktiv und gezielt zur Steuerung von Programmen beizutragen und deren „innerer Kompass“ zu sein. Beide sind zwar essenzielle Bestandteile der EZ, bleiben jedoch in vielen, vor allem kleineren Organisationen „im Tauziehen zwischen Rechtfertigung und institutionellem Lernen“<sup>7</sup> hinter ihren Möglichkeiten zurück. Gleichwohl haben der interne und externe Druck auf die EZ insgesamt zu einer Professionalisierung von M&E geführt: Die Bundeshaushaltsordnung verlangt eine Zielerreichungs-, Wirkungs- und Wirtschaftlichkeitskontrolle, und fast jede Nichtregierungsorganisation (NRO), politische Stiftung

oder andere Organisationen verfügen inzwischen über eigene M&E-Einheiten und -Strukturen. Neben der inhaltlichen Ausrichtung sind es vor allem Finanz- und Verwaltungsvorgänge, die im Vordergrund der Rechtfertigungsdebatte stehen. Moderne Verwaltungssysteme erfassen Finanzflüsse jedoch schnell und präzise, und die Gefahr, dass Gelder nicht zweckgebunden verausgabt werden, ist reduziert worden. Es bleiben Vorbehalte gegenüber der Wirksamkeit von Entwicklungszusammenarbeit, denen weniger deutlich begegnet werden kann. Dies auch, weil Wirkungsmessung im Verhältnis zum Finanzmonitoring vor allem im Governance-Bereich als weniger eindeutig und realisierbar wahrgenommen wird, mithin kaum als Orientierungshilfe dient. So finden die Ergebnisse eher selten Eingang in Strategieprozesse oder politische Diskussionen und sind entsprechend schwerer abrufbar.

### **Die größte Herausforderung liegt im Verständnis des Wirkungsbegriffes.**

---

Die Gründe hierfür sind vielfältig. Einerseits werden M&E-Mechanismen durch ausführende Akteure der EZ häufig auf den Legitimationsaspekt reduziert und ihre Chancen und Möglichkeiten (etwa zur Selbstreflexion und zum Anstoßen von Lern- und Strategieprozessen) nicht ausreichend wahrgenommen. Andererseits wirken eine allzu fachspezifische Begrifflichkeit und Methodik oftmals technisch und wenig attraktiv, für Governance-Anliegen meist sperrig und inadäquat. Entsprechend zeigt sich auch die Kommunikation der Ergebnisse nur wenig eingängig und nicht immer auf die Adressaten hin maßgeschneidert, somit kaum ansprechend oder auch nur verständlich.

Doch die größte Herausforderung liegt im Verständnis dessen, was erwiesen werden soll, des Begriffs also, den es zu dokumentieren gilt: der „Wirkung“. Denn was ist „Wirkung“ und welche Definition liegt ihr zugrunde?

## Wirkung – zwischen Output, Outcome und Impact

Die erneuerte Diskussion um den Wirkungsbegriff entspricht einer grundsätzlichen Veränderung, welche die deutsche Entwicklungszusammenarbeit in den vergangenen beiden Jahrzehnten erlebt hat und die sich beispielsweise in der Gründung der größten und bekanntesten EZ-Institution in Deutschland veranschaulichen lässt: 2011 entstand die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) durch die Fusion der Deutschen Gesellschaft für Technische Zusammenarbeit (GTZ), der Internationalen Weiterbildung und Entwicklung gGmbH (InWEnt) und des Deutschen Entwicklungsdienstes (DED). Schon vor der Umfirmierung begann ein Trend, für den nicht nur die GIZ steht: die Tendenz, technische Zusammenarbeit, Aufbau von Infrastruktur bis hin zu Nothilfe in übergeordnete gesellschaftliche und politische Ebenen und Strukturen einzubetten und etwa durch Maßnahmen zur „Guten Regierungsführung“, Unterstützung in Verwaltungsreformen und Politikberatung zu flankieren. Allein im Bereich „Good Governance“ konnte die GIZ nach eigenen Angaben seit 2008 einen 65-prozentigen Zuwachs an Projekten verzeichnen, finanziell hat sich das Volumen im gleichen Zeitraum mehr als verdreifacht.<sup>8</sup>

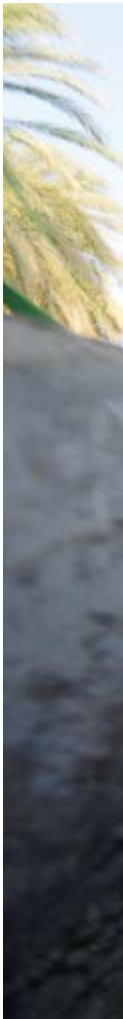
Aus TZ wurde also IZ (Internationale Zusammenarbeit), was die Frage nach Wirkung und deren Messbarkeit erneut und anders stellte. Dies auch vor dem Hintergrund, dass sich die internationale Debatte geändert hat und – in der Folge – entsprechende Abkommen und internationale Verpflichtungen, die auch Deutschland eingegangen ist, dem Wirkungsaspekt eine zentralere Rolle eingeräumt haben.<sup>9</sup> Mag es bei technischen Maßnahmen noch gelingen, zumindest vordergründig eine Kausalität zwischen einer Handlung und entsprechenden Veränderungen herzustellen, wird dies umso schwieriger, je komplexer sich die Wirkungszusammenhänge darstellen. Ein neuer Brunnen etwa ermöglicht die Bewässerung von Ackerflächen, diese liefern reichere Ernten und führen zu einer verbesserten Ernährungssicherheit in einer Region. Bei einer Zusammenarbeit jedoch, die auf längerfristige

Veränderungen etwa von Verhaltensmustern ausgelegt ist, ist es weitaus schwieriger, Kausalitäten zu belegen oder auch nur zu erkennen: Haben Beratungsleistungen in kommunaler Selbstverwaltung und Workshops mit Beamten dazu geführt, dass eine verbesserte Budgetplanung in der Kommune vorangetrieben wird? Sorgt dies tatsächlich dafür, dass Korruption zuerst auf kommunaler Ebene und später auf nationaler Ebene sinkt?

## Bei politischer Bildung oder Dialogprogrammen ist Wirkung deutlich schwerer messbar.

Die Flankierung technischer Zusammenarbeit durch Maßnahmen im politischen Bereich und der öffentlichen Verwaltung hat die Aufgabe von Evaluatoren und M&E-Experten zunächst kompliziert. Wo Interventionen durch besagte Brunnen oder verteilte Saatgutbeutel konkrete und direkte Auswirkungen auf die Betroffenen hatten und dies auch durch Beobachtungen oder klassische Quantifizierung belegt werden kann, sind komplexe strukturelle Veränderungen deutlich schwerer messbar zu machen und deren Wirkungsbeschreibung entsprechend breit. Geht man gar auf die Ebene politischer Bildung, politischer Dialogprogramme oder ähnlicher Maßnahmen über, so wird die Definition und Messbarkeit von Wirkung noch schwieriger.

Daher verwundert es nicht, dass sich die Diskussion um Wirksamkeit in der EZ aufgefächert hat und differenzierter, aber auch diffuser zu werden scheint. International agierende Organisationen legen – je nach Mandat – unterschiedliche Maßstäbe in der Bewertung ihrer Bestrebungen an. So gibt sich die Weltbank – als Entwicklungsbank – beispielsweise engere Kriterien und definiert Wirkung als „Interessensgegenstand mit oder ohne Intervention:  $Y_1 - Y_0$ “<sup>10</sup>. Wirkung wird danach nur erzielt, wenn der Unterschied, den eine gewisse Intervention ausgelöst hat, zweifelsfrei gemessen werden kann, was sich allerdings





nicht ohne den aufwändigen Einbezug von Kontrollgruppen realisieren lässt.

Der Entwicklungshilfesausschuss der OECD (OECD DAC) hingegen definiert Wirkung als „positive und negative, kurz- oder längerfristige Effekte, die indirekt oder direkt und intendiert oder unintendiert durch Entwicklungsinitiativen hervorgerufen werden“<sup>11</sup>. Diese zweite Definition, die auch die Grundlage für das Verständnis der meisten deutschen EZ-Organisationen bildet, ist breiter gefasst, was für eine multilaterale Organisation wie die OECD (Organisation für

wirtschaftliche Zusammenarbeit und Entwicklung) auch nachvollziehbar und angemessen ist. Man geht davon aus, dass Wirkung auf verschiedenen Ebenen stattfinden kann. Im Englischen wird dies alleine schon sprachlich sichtbar: Man unterscheidet grundsätzlich zwischen Output, Outcome und Impact. Während ein Output ein üblicherweise einfach zu messendes direktes Ergebnis einer Intervention, beispielsweise eines Workshops, ist, sind die beiden letzten Kategorien schwerer greifbar. Outcomes sollten Justierungen im Verhalten bestimmter Zielgruppen beinhalten, also Effekte auf der Zielgruppenebene erreichen,



Selten so offensichtlich: Ist die Wirkung einer neuen Bewässerungsanlage noch relativ gut messbar, so lassen sich Kausalitäten bei komplexen strukturellen Veränderungen in der Gesellschaft weit schlechter belegen.

Quelle: © Amr Abdallah Dalsh, Reuters.

oder aber auch Zustandsveränderungen bewirken (etwa, dass die Mangelernährung in einer Zielpopulation um einen „Faktor X“ reduziert worden ist), während die Impact-Ebene darüber hinausgeht und eine Makroebene, wie die Gesellschaft oder einige ihrer Teile, umfasst.<sup>12</sup> Die Einflusssebenen, auf denen sich Output, Outcome und Impact bewegen, werden entsprechend als „sphere of control“ (Output), „sphere of influence“ (Outcome) und „sphere of interest“ (Impact)<sup>13</sup> beschrieben. Vor dem Hintergrund dieser Differenzierungsoptionen hat sich das Englische als Lingua franca auch in der deutschen EZ durchgesetzt und bietet die Möglichkeit genauerer Unterscheidung der verschiedenen Wirkungsebenen und somit der Veränderungen, die erzielt worden sind.

## Rechtfertigungsdruck kann die Wahrnehmung von M&E-Einheiten als Kontrollgremium verstärken.

---

Vor dem Hintergrund der Schwierigkeiten in der Begriffsbestimmung und Messbarkeit des Erreichten verwundert es nicht, dass Rechtfertigung und Nachweis einer „Sinnhaftigkeit“ im Sinne einer „Wirksamkeit“ von Projekten eine Herausforderung bedeuten. Zumal die Terminologie auch nicht von allen EZ-Akteuren gleichermaßen und einheitlich nachvollzogen wird, was die Kommunikation mit Zielgruppen in Deutschland erschwert. Schon die Begriffsbestimmung stellt eine Hürde dar – und weitere Hürden kommen hinzu.

### Problematische Wahrnehmungen

Der erhöhte Rechtfertigungsdruck der vergangenen Jahre sowie die immer komplexer werdende Begriffsbestimmung haben dazu geführt, dass M&E-Einheiten in nahezu allen Organisationen der deutschen Entwicklungszusammenarbeit professionalisiert und personell aufgestockt wurden. Auch die Nachfrage nach Aufbaustudiengängen in diesem Bereich ist gestiegen und zahlreiche

Beratungsfirmen konzentrieren sich ausschließlich auf Evaluationen für EZ-Einheiten weltweit.

Die Professionalisierung der M&E-Einheiten sorgt einerseits dafür, dass Daten rascher und verlässlicher aggregiert werden können und die Voraussetzungen für eine kompetente und schnelle Beantwortung diverser Anfragen, beispielsweise durch Ministerien oder Fraktionen im Bundestag, besser erfüllt werden können. Der Rechtfertigungsdruck kann unter M&E-Einheiten andererseits aber auch zu einer Selbst- und Fremdwahrnehmung als Kontrollgremium führen. So verdrängt die Dringlichkeit von Wirkungskontrolle im Tagesgeschäft mitunter einige Vorzüge von M&E-Maßnahmen, die auch darüber hinaus durchaus nutzbar gemacht werden können: etwa als Unterstützung bei strategischen Entscheidungen, dem Aufzeigen von Nischen, in denen sich Alleinstellungsmerkmale konkurrierenden Institutionen gegenüber herausarbeiten ließen, und vieles mehr. Gleichzeitig spielen sie auch eine essenzielle Rolle im institutionellen Lernprozess. Auch wenn Projekte regional und themenspezifisch unterschiedlich aufgezogen werden, so gibt es unter Umständen doch Ähnlichkeiten in den Abläufen. Eine Einheit, die Einblicke in sämtliche Projekte weltweit hat, kann für den institutionellen Lernprozess von größtem Wert sein, wird aber oftmals nicht entsprechend in Anspruch genommen.

### Theorie vs. Praxis

Monitoring und Evaluation geschehen auf allen Ebenen in der Entwicklungszusammenarbeit: auf der Mikroebene (Maßnahme), Mesoebene (Projekt) und Makroebene (Programm oder Sektorprogramm). So mag etwa ein Workshop bereits direkten Einfluss auf die Teilnehmer haben, der auch nachgewiesen werden kann. Dass diese Teilnehmer dann aber auch ihr Umfeld beeinflussen und dies wiederum dazu führt, dass ein größerer, übergeordneter Bereich sein Verhalten ändert, kann zwar angenommen, aber weniger einfach nachgewiesen werden. Dies sind zwar verschiedene Aspekte mit unterschiedlichen Ansprüchen an M&E. Sie leiten sich aber von der untersten Ebene her und hängen miteinander

zusammen: Sind also Daten auf der Maßnahmenebene nicht korrekt erfasst worden, wird es schwer, ein möglichst adäquates Bild auf einer übergeordneten Ebene zu erhalten, und umso schwieriger, je weiter man schaut.

## **Vor Ort erfasste Daten entsprechen oft nicht den Standards, die von M&E-Einheiten erwartet werden.**

Ein ausgeklügeltes Monitoring und sauber erfasste Daten auf der Mikroebene sind eine unabdingbare Voraussetzung und das Grundgerüst für alle weiteren Datenerhebungen. Ein vorab durchdachtes System mit Wegmarken und Möglichkeiten zur Umsteuerung und Intervention ist unerlässlich, benötigt aber sowohl personelle Ressourcen im Feld als auch finanzielle Ressourcen – sowie bei allen Beteiligten ein klares Verständnis des zu beobachtenden Gegenstands. Oft reichen die oben genannten Ressourcen jedoch nicht aus, sodass es nicht überrascht, dass erfasste Daten vor Ort häufig nicht den Standards entsprechen, die von M&E-Einheiten erwartet werden.<sup>14</sup> Fragebögen, in denen die Zufriedenheit der Teilnehmer oder deren Eindrücke abgefragt werden, tragen meist wenig zur Wirkungsmessung bei. Zudem ist das Monitoring komplexer struktureller Veränderungen oder der Wirksamkeit von sogenannten Track-2-Formaten (informelle Gesprächskanäle, häufig in diplomatisch oder politisch heiklen Kontexten) sowie von Netzwerken häufig sehr kompliziert und benötigt längerfristige Abfragen bei Teilnehmern und Beobachtungen der Projektmanager – zwei Umstände, die nur schwer nachzuhalten sind. Auch entfernen sich die zumeist auf dem Papier festgelegten Wirkungsindikatoren zuweilen vom Projektverlauf oder der Realität. Dass dieser Umstand eine Evaluation auf Programmebene erschwert, liegt nahe. Evaluatoren müssen oftmals mit anekdotischen Evidenzen hantieren, sich auf intuitive Aussagen von Projektverantwortlichen und deren Vorgesetzten verlassen und können Erkenntnisse kaum quantifizieren. Hinzu kommt das althergebrachte

Gefühl, dass M&E überwiegend der Kontrolle und Bewertung dient. So fällt es schwer, offen mit ausbleibenden oder unerwünschten Ergebnissen umzugehen, selbst dann, wenn es ohne eigenes Zutun und unverschuldet geschieht. Auch dies behindert M&E-Maßnahmen in ihrer Akzeptanz und in ihren Gestaltungsmöglichkeiten.

## **Evaluierungen in der Praxis – Konsequenzen für Projekte?**

Wie häufig ist der offene Umgang mit negativen Ergebnissen in der Praxis? Bilden Evaluierungen es auch ab, wenn Projekte nicht erfolgreich sind? Und wie steht es um die Urteilskraft von Evaluierungen, wie nah kommen sie der Realität vor Ort? Dies sind berechtigte Fragen. Das beschränkte Verständnis der Rahmenbedingungen (und somit des Projekts) ist ein Vorwurf, den sich Evaluatoren gelegentlich gefallen lassen müssen: Eigene Daten werden in einem begrenzten Zeitraum erhoben und Daten, die vorher gesammelt wurden, lediglich analysiert. Von daher sind Evaluationen Bewertungen und Beratungen „nur“ anhand von (wenngleich fundierten) Annahmen, die – je nach Datenlage – zwar nah an die Projektrealität heranreichen, diese indes nicht vollständig abbilden. Ein Anspruch, den Evaluierungen aber auch nicht haben, weshalb sie selbst eher selten zu „Radikalempfehlungen“ tendieren und auch von außen nur als eines von vielen Elementen in Steuerungs- und Strategieprozessen wahrgenommen werden sollten. Dennoch führt die Angst, dass ein Gutachten, das negativ ausfällt, über die Zukunft des Projektes und des eigenen Werdegangs entscheidet, zu Abwehrreflexen bei den Projektverantwortlichen (und beim Gutachter möglicherweise zur Befürchtung, keine Folgeaufträge zu bekommen).

Evaluierungen mit kritischem Ausgang sind dennoch nicht selten, wenngleich der Blick meist konstruktiv ist und der Fokus auf den Errungenschaften liegt. Selbst in einem Land wie Afghanistan, in dem trotz erheblicher Anstrengungen im *Peace-* und *State-Building*-Bereich durch verschiedene Akteure die volatile Sicherheitslage und ein Wiedererstarken von Gruppierungen wie den Taliban oder lokalen Zellen des sogenannten

Islamischen Staates etwaige Erfolge immer wieder konterkarieren, sind Evaluationen (etwa der Bundesregierung) zu dem Urteil gekommen, dass „die wirtschaftliche, soziale und gesellschaftliche Entwicklung Afghanistans [...] seit dem Sturz der Taliban bereits bemerkenswerte Fortschritte gemacht“ und „Deutschland [...] in den vergangenen 18 Jahren – gerade auch mit seiner Entwicklungszusammenarbeit – zu dieser Entwicklung beigetragen und gemeinsam mit der afghanischen Regierung sowie der internationalen Gemeinschaft wichtige Grundlagen für den sozialen und wirtschaftlichen Fortschritt des Landes geschaffen“ hat.<sup>15</sup> Dies ist genauso richtig wie die Feststellung, dass grundsätzlich „der politische Wille, die politische Durchsetzungsfähigkeit, die politischen Wertvorstellungen sowie die Gestaltung der Wirtschaftsordnung im Partnerland“<sup>16</sup> die Arbeit behindern oder erleichtern.

### Umgang mit Misserfolgen und Fehlern

Dass also vor allem Rahmenbedingungen über Erfolg und Misserfolg von EZ-Projekten bestimmen, wird immer wieder zu beobachten sein. Sie (und weniger Evaluierungen) entscheiden – neben taktischen Gründen – am häufigsten über einen Fortgang der Projekte oder überhaupt den Verbleib im Partnerland. Die unüberschaubare Lage im Jemen etwa, nach dem Sturz des damaligen Präsidenten Ali Abdullah Saleh, zwang die meisten internationalen EZ-Organisationen, darunter auch die GIZ, ihre entsandten Kräfte abzuziehen und in „Fernsteuerung“ gemeinsam mit lokalen Kräften zu arbeiten. Dass die Zusammenarbeit mit dem Jemen den verschärften Bedingungen seitdem angepasst werden musste, liegt auf der Hand.<sup>17</sup> Allerdings stellt dies nicht zwangsläufig die Wirksamkeit der Maßnahmen infrage, sofern die Anpassung – und hierzu können Evaluationen dienlich sein – zweckmäßig und klug erfolgt ist. Am Ende liegt es jedoch im Ermessen der Organisation, ob und wie ein Projekt fortgeführt wird. Ebenso verhält es sich beim Umgang mit Misserfolgen und Fehlern: Die Frage, inwiefern Kritik und Anregungen aus Evaluierungen etwa in Folgeprojekten ausreichend berücksichtigt werden, kann pauschal nicht beantwortet werden. Es ist die Organisation, die – unter Berücksichtigung



aller genannten und relevanten Aspekte – letztlich hierüber entscheidet.

Dies gilt auch für den Umgang mit gewonnenen Daten und Erkenntnissen, die hierauf beruhen. Daten sind trockene Zahlen und Fakten. So ist es





Schwer getroffen: Wenn sich in einem Partnerland wie im Jemen die Rahmenbedingungen massiv verschlechtern, dann bleibt das nicht ohne Folgen für die Entwicklungszusammenarbeit. [Quelle: © Abduljabbar Zeyad, Reuters.](#)

entscheidend, dass die gesammelten Informationen aufgearbeitet werden, und zwar zielgruppenadäquat.<sup>18</sup> Meist geht dies jedoch nicht über den Evaluationsbericht hinaus. Dieser dient oft als alleinige Grundlage für jede Art von Kommunikation – sei es gegenüber der Geschäftsleitung,

in der Verbindung zum Zuwendungsgeber oder in der Bearbeitung von Anfragen im Rahmen des Informationsfreiheitsgesetzes. Deren mitunter sehr unterschiedliche Anforderungen und Interessen müssen in der Kommunikation berücksichtigt werden, um überhaupt „gehört“ zu werden.<sup>19</sup>

Hier herrscht vor allem auf Seiten der Evaluierungseinheiten eine Bringschuld gegenüber den verschiedenen Akteuren. Sie müssen in der Lage sein, die gelieferten Ergebnisse für die Empfänger so zu interpretieren und aufzubereiten, dass die Nutzbarkeit von Evaluationen – und letztlich der Nutzen der Projektarbeit selbst – daraus ersichtlich ist. Schließlich sind sie das Bindeglied zwischen der Organisation und den – zumeist externen – Evaluatoren und können am ehesten einschätzen, welche Wege beschritten werden sollten – und welche nicht. Zudem können sie auch intern nur so Lernprozesse anstoßen oder erwirken, in die übergeordnete Strategieberatung eingebunden zu sein.

### Eine konstruktive Fehlerkultur würde dazu beitragen, die gewonnenen Erkenntnisse weltweit anderen Projekten zugutekommen zu lassen.

---

#### Imagewandel – von Kontrolle zu Beratung

Es bedarf einiger Schritte, um die beschriebenen Herausforderungen zu bewältigen. Zunächst muss in der **Wahrnehmung** ein Umdenken stattfinden. M&E-Einheiten müssen den Spagat schaffen, ihre Kontrollfunktion wahrzunehmen, ohne „Hüter der Indikatoren“ zu sein. Vielmehr müssen sie (auch) als Berater auftreten und Lösungen anbieten und nicht nur wertend zurück, sondern konstruktiv nach vorne schauen. Ein zugewandtes Auftreten gehört zu dem Anspruch, „Freund und Helfer“ zu sein. Umgekehrt ist auch das Evaluertwerden mit Scheu verbunden, schwingt eine Bewertung doch immer mit. Nur mit einer offenen Fehlerkultur und dem Willen zum Lernen und zur Veränderung wird es gelingen, Evaluationen als hilfreiches Instrument zu betrachten und gewinnbringend einzusetzen. Der Rechtfertigungsaspekt sollte dem Lernprozess dabei nicht im Wege stehen: „Evaluierung dient hauptsächlich zwei Zwecken: Rechenschaftsablegung und Lernen. Akteure der

Entwicklungszusammenarbeit haben sich bisher in der Regel auf Ersteres konzentriert und die Verantwortung hierfür an zentralisierte Einheiten übertragen. Aber Evaluierung zum Zwecke des Lernens ist das Gebiet, für das Beobachter für die Gegenwart und Zukunft den größten Bedarf ausmachen.“<sup>20</sup> Insbesondere sollten die Ergebnisse aus Monitoring- und Evaluationsprozessen dazu dienen, Projekte weiterzuentwickeln und in ähnlichen Kontexten bereits entstandene Mängel zu vermeiden. Die Bedeutung der Lernerfahrung aus M&E und die gezielte Nutzbarmachung dieses Aspekts wurden auch im letzten OECD Peer Review angemerkt.<sup>21</sup>

Die **Diskrepanzen zwischen Theorie und Praxis** lassen sich nur schwer überwinden, da Professionalität im Feld nicht immer zu gleichen Teilen erwartet oder gelebt werden kann. Nicht immer ist es möglich, ein adäquates, den Anforderungen entsprechendes Monitoring zu gewährleisten, da vor allem personelle Ressourcen hierfür nicht selten fehlen. Selbst wenn Daten erhoben und ausgewertet werden können (was vor dem Hintergrund vielfältiger und anspruchsvoller Aufgaben von Projektverantwortlichen nicht immer gelingt), so ist deren Analyse und Aufbereitung meist doch eine Herausforderung: Mittel- und längerfristige Wirkungen auf politischer oder ökonomischer Ebene werden ad hoc nicht beurteilt werden können. Eine Möglichkeit wäre vor allem für kleinere Organisationen, die ohne „Monitoring-Beauftragten“ auskommen müssen, Stichproben hervorzuheben und statt mit einer Quantifizierung mit qualitativen Datenerhebungsmethoden zu arbeiten. Diese mögen zwar leichter manipulierbar sein, sind aber aussagekräftiger als nicht durchgängig korrekt erhobene Daten. Dabei sollte weniger die Angst vor dem Eingeständnis fehlender Wirkungen im Vordergrund stehen als der positive Lerneffekt. Eine konstruktive Fehlerkultur würde zudem dazu beitragen, aus Fehlern gewonnene Erkenntnisse auch weltweit anderen Projekten zugutekommen zu lassen, vorausgesetzt, die Einsichten werden erkannt, thematisiert und kommuniziert. Abgesehen davon kann es auch sein, dass Umstände, die außerhalb des eigenen Einflussgebietes liegen (etwa ausbrechende Konflikte oder Naturkatastrophen), Vorhaben torpediert und unwirksam

gemacht haben. Auch hier kann der Umgang mit Fehlern, sofern kommuniziert, Anregung und Hilfe, jedoch kein Zwang zum Umsteuern in Folgeprojekten sein.

**Kommunikation** ist also eine wichtige Voraussetzung für Lernprozesse und der Schlüssel, wenn es darum geht, die Relevanz von M&E-Maßnahmen auf Mikro-, Meso- und Makroebene zu steigern. Und nicht zuletzt, um die Ergebnisse zu dokumentieren und weiterzugeben: Wer ist der Empfänger der entsprechenden Informationen und was erwartet er? Wie methodisch bewandert ist er und welche zeitlichen Ressourcen stehen zur Verfügung? Es gilt, die Datenaufbereitung und Kommunikation zielgruppengerecht passend zu machen; schließlich sind für eine parlamentarische Anfrage oder ODA-Statistiken andere Daten und andere Arten der Datenaufbereitung notwendig als für die Kommunikation von Wirkungen an die Medien oder die Öffentlichkeit.

Wie eine Unternehmensberatung sollte sich auch jede M&E-Einheit diese und ähnliche Fragen stellen, bevor sie Ergebnisse von Evaluationen vermittelt. Insbesondere in strategischen und politischen Fragen muss sie potenzielle Auswirkungen bereits auf ein politisches Level extrapolieren, aufbereiten und vorwegnehmen können, damit Transferleistungen anderer Akteure vorgegriffen werden kann. Gerade im Politik- und Governance-Bereich sollte sie aber auch Vorsicht walten lassen und sich fragen, welche Aspekte sie vielleicht bewusst auch nicht kommuniziert: Mit Blick auf undemokratisch und autoritär verfasste Gesellschaften und Strukturen könnte ein Zuviel an Transparenz in einigen Fällen kontraproduktiv oder gar gefährlich sein. Von daher sollte hier eher die Maxime gelten, dass normative Ansprüche offen kommuniziert, operative Informationen (etwa mit Blick auf Quellenschutz) aber vertraulich behandelt werden sollten.

## Fazit und Ausblick

Die „richtige“ Kommunikation wird jedenfalls wichtiger, denn der Trend der vergangenen Jahre wird sich verstärken: Staatliche Ausgaben,

insbesondere in der Entwicklungszusammenarbeit, werden vermehrt von politischen Akteuren und Teilen der Gesellschaft kritisch begleitet. Allen voran die Coronapandemie und ihre Auswirkungen auf die deutsche Wirtschaft werden die Diskussionen um EZ-Ausgaben weiter verschärfen. Entsprechend wird der Druck auf EZ-Akteure wachsen, die Wirksamkeit ihrer Maßnahmen zu dokumentieren und zu kommunizieren – wenn nicht sogar zu vermarkten. Von Herausforderungen durch Systemkonkurrenten wie China, die Golfstaaten oder auch die Türkei gar nicht erst zu reden, die alle Mittel nutzen, um nicht nur wirtschaftlich Einfluss zu gewinnen. Von daher sind überzeugende Konzepte und dokumentierte Wirkungen nicht zu vernachlässigende „Selling Points“ für deutsche EZ-Organisationen.

Auch das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) hat dies erkannt und Themen wie Wirksamkeit und Datenverfügbarkeit im Rahmen der „BMZ 2030“-Reform auf seine Agenda gesetzt.<sup>22</sup> Dies zeigt, dass das BMZ sich des zunehmenden Rechtfertigungsdrucks deutlich bewusst ist. Somit wird der Wirkungsbegriff – und damit auch die Wirkungsdokumentation und -kommunikation – an Relevanz noch zunehmen und immer weniger ein Nischenthema für interne Prozessbetrachtungen sein. Diese Entwicklung gilt es vorwegzunehmen und sich vorausschauend und proaktiv aufeinander zuzubewegen: Nur so können M&E-Experten und politische Entscheidungsträger ein strategisch nutzbares Monitoring und Evaluieren zum internen Kompass der EZ weiterentwickeln, welcher ihr Schiff – auch durch unruhige Gewässer – weiter in Richtung Erfolg und Akzeptanz navigiert.

---

**Dr. Angelika Klein** ist Leiterin der Stabsstelle Evaluierung der Konrad-Adenauer-Stiftung.

**Lukas Kupfernagel** ist Referent in der Stabsstelle Evaluierung der Konrad-Adenauer-Stiftung.

- 1 Staatssekretär Martin Jäger bei der Vorstellung des Evaluierungsberichts 2020 der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ), 25.02.2021.
- 2 Easterly, William 2001: *The Elusive Quest For Growth: Economists' Adventures and Misadventures in the Tropics*, Cambridge, S. xii–xiii.
- 3 Biekart, Kees / Verkoren, Willemijn / Kamphuis, Bertine 2013: *State Building in a Rentier State: How Development Policies Fail to Promote Democracy in Afghanistan*, in: *Development and Change*, 44: 3, S. 501–526.
- 4 Mason, Jeff 2020: *Exclusive: Trump proposes 21 % cut in U.S. foreign aid in budget proposal*, Reuters, 09.02.2020, in: <https://reut.rs/3ftUQZY> [08.03.2021].
- 5 Gleser, Solveig H. / Schneider, Sebastian H. 2018: *Meinungsmonitor Entwicklungspolitik 2018: Einstellungen zu Entwicklungszusammenarbeit und nachhaltiger Entwicklung*, Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), in: <https://bit.ly/3ypvjcU> [17.05.2021].
- 6 Quirk, Patrick 2014: *(Re)Emerging Aid Donors in the Reshaping World Order*, The German Marshall Fund of the United States (GMF), 11.07.2014, in: <https://bit.ly/33XAgvG> [17.05.2021].
- 7 Guijt, Irene 2010: *Accountability and Learning: Exploding the Myth of Incompatibility between Accountability and Learning*, in: Ubels, Jan / Acquaye-Baddoo, Naa-Aku / Fowler, Alan (Hrsg.): *Capacity Development in Practice*, Earthscan, S. 339, in: <https://bit.ly/3hxwSzE> [19.05.2021].
- 8 GIZ 2021: *Evaluierungsbericht 2020 – Wissen nutzen*, 10/2020, in: <https://bit.ly/3v31aOr> [06.03.2021].
- 9 Insbesondere die Erklärung von Paris über die Wirksamkeit der Entwicklungszusammenarbeit 2005 u.v.m.
- 10 Hearn, Simon / Buffardi, Anne 2016: *What is Impact?*, Overseas Development Institute (ODI), 02/2016, S. 7, in: <https://bit.ly/3hC8ZH3> [19.05.2021].
- 11 Ebd.
- 12 Mercator Stiftung 2016: *Leitfaden Wirkungsmodell*, 09/2016, in: <https://bit.ly/3bG34wO> [08.03.2021].
- 13 Hearn, Simon 2008: *Outcome Mapping*, ODI, in: <https://bit.ly/3bBXjjZ> [04.05.2021].
- 14 Pasanen, Tiina 2017: *Monitoring and evaluation: five reality checks for adaptive management*, ODI, 14.12.2017, in: <https://bit.ly/3bVGjW7> [10.04.2021].
- 15 Deutscher Bundestag 2019: *Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Markus Frohnmaier, Dietmar Friedhoff, Ulrich Oehme und der Fraktion der AfD, Entwicklungszusammenarbeit mit Afghanistan*, Drucksache 19/8231, 07.03.2019, in: <https://bit.ly/3osUSoW> [06.05.2021].
- 16 GIZ 2021, N. 8, S. 13.
- 17 Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) 2021: *Jemen – Land in humanitärer Krise*, in: <https://bmz.de/de/laender/jemen> [06.05.2021].
- 18 Dabei muss der Begriff der Zielgruppen von jenem der Stakeholder unterschieden werden, um den Unterschied zwischen jenen Gruppen, bei denen Wirkungen entstehen (Stakeholder) und jenen, an die Wirkungen kommuniziert werden (Zielgruppen), deutlich zu machen.
- 19 Grünhaus, Christian / Rauscher, Olivia 2021: *Impact und Wirkungsanalyse in Nonprofit Organisationen, Unternehmen und Organisationen mit gesellschaftlichem Mehrwert*, Wirtschaftsuniversität Wien, 02/2021, in: <https://bit.ly/3wbWXrY> [17.05.2021].
- 20 Serat, Olivier 2009: *Learning from Evaluation*, Knowledge Solutions 44, Asian Development Bank (ADB), 05/2009, S. 1, in: <https://bit.ly/3v5kkmX> [08.03.2021].
- 21 Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) 2013: *Evaluating Development Activities, 12 Lessons from the OECD DAC*, in: <https://bit.ly/3wlpvV9> [17.05.2021].
- 22 BMZ 2018: *Entwicklungspolitik 2030, Neue Herausforderungen – neue Antworten*, 10/2018, in: <https://bit.ly/33XcbvW> [10.04.2021].