

Herausforderungen und Grenzen, Ansätze und Perspektiven der Evaluation in der Entwicklungszusammenarbeit¹

Reinhard Stockmann, Universität des Saarlandes

1. Diskrepanz in der Beurteilung des entwicklungspolitischen Erfolgs

Bei der Beurteilung entwicklungspolitischer Wirkungen von Förderprogrammen der Entwicklungszusammenarbeit (EZ) fällt vor allem die große Diskrepanz zwischen Fremd- und Selbsteinschätzung auf. Während sich die staatlichen und nicht-staatlichen Fördermittelgeber mit Erfolgsmeldungen geradezu überbieten, kann die Kritik der Gegner der EZ kaum fundamentaler sein. Die Hilfe nutze vor allem den herrschenden Eliten, stabilisiere verkrustete und ungerechte Machtstrukturen, verhindere politische und institutionelle Reformen (Myrdal 1984, Elsenhans 1984 u.v.a.), störe die freie Entfaltung der Marktkräfte (Bauer 1984), diene der Ausbeutung der Dritten Welt und ermögliche „eine neue Form der Kontrolle durch Heerscharen von Entwicklungsplanern, Finanzexperten und Regierungsberatern“ (Simon 1991), schade schließlich allen, „denen sie angeblich nützen soll“, stifte „nichts als Unheil“ oder wirke gar „tödlich“, so dass schließlich der Schluss gerechtfertigt erscheint: „Ohne Entwicklungshilfe ginge es den Menschen in den Ländern der Dritten Welt besser“ (Erler 1985). Diese Kritik ist so alt wie die Entwicklungshilfe selbst und zieht sich bis heute (vgl. z.B. neuerdings Bode 2002) als Begleitkanon durch die Gazetten und prägt die öffentliche Meinung.

Ganz anders tönen die Wirkungsbeurteilungen der Durchführungsorganisationen. In dem kürzlich veröffentlichten 5. Ergebnisbericht der Kreditanstalt für Wiederaufbau², die in Deutschland für die Abwicklung der Finanziellen Zusammenarbeit (vornehmlich Kredite) zuständig ist, heißt es: „Drei Viertel der untersuchten Vorhaben sind entwicklungspolitisch erfolgreich“ (KfW 1999: 4). Um jeden Zweifel an der Glaubwürdigkeit der Ergebnisse zu zerstreuen, wird vorsorglich bekannt: „Die KfW hat weder Anlass noch Interesse, die Erfolgsbilanz ‚schön zu färben‘“ (KfW 1999: 3). Selbstverständlich kann die Deutsche Gesellschaft für Technische

1 Dem Artikel liegt ein Vortrag auf der Jahrestagung der SEVAL am 7. Juni 2002 in Bern zugrunde.

2 Die KfW fördert Investitionen durch Kredite. Auftragsvolumen ca. 2 Milliarden Euro/Jahr.

Zusammenarbeit³ ähnlich gute Ergebnisse vorweisen. Da wollen die Nicht-Regierungsorganisationen (NRO) nicht zurückstehen und übertreffen mit ihren Erfolgsmeldungen sogar noch die der staatlichen Träger.

Wie passen diese grandiosen Erfolgsbilanzen zu den in den Medien regelmäßig verbreiteten Katastrophenmeldungen über fehlgeschlagene Projekte, unsensible Eingriffe in gewachsene Strukturen, entwicklungspolitisch irrelevante Programme und die wachsende Armut in der Dritten Welt?

Könnte es sein, dass die EZ doch weitaus besser ist, als ihr Ruf oder liegt es an der subjektiven Perspektive der Akteure? Auf welche Daten und Informationen stützen eigentlich die Kritiker und Befürworter ihre Positionen? Mit welchen Methoden werden sie erhoben und sind sie überhaupt geeignet, die Wirkungen entwicklungspolitischer Massnahmen zu erfassen? Dies sind wichtige Fragen an die Evaluation der EZ, die hier thematisiert werden sollen, um anschliessend die Aufgaben und Perspektiven der Evaluation für die Zukunft entwickeln zu können.

2. Quellen entwicklungspolitischer Erfolgsbeurteilung

Festzuhalten ist zunächst, dass eine Vielzahl der Kritiker der EZ nicht auf die Ergebnisse fundierter Evaluationen zurückgreifen, sondern dass sie sich zumeist auf eigene Erfahrungen und Beobachtungen stützen. Durchweg handelt es sich um Betroffenheitsberichte, in denen Einzelbeobachtungen unsystematisch aneinander gereiht werden, die Stichprobe der ausgewählten Fälle sehr klein ist, die Beurteilungskriterien nicht offen gelegt werden, differenzierte Untersuchungsmethoden nicht angewandt werden, kaum zwischen kurz- und langfristigen, intendierten und nicht-intendierten Wirkungen unterschieden wird, die Ursachenanalyse an der Oberfläche verbleibt und dennoch weitreichende Schlussfolgerungen gezogen werden.

Das Gegenstück zu diesen individualistischen, meist emotional eingefärbten Fallbetrachtungen stellen Untersuchungen mit makrobezogenen Entwicklungsdaten dar. Hierfür werden in der Regel aus Aggregatdaten gewonnene ökonomische und gesellschaftliche Globalindikatoren (wie z.B. das BIP, Analphabetismusrate, Säuglingssterblichkeit etc.) verwendet. Zumeist wird im Rahmen von Regressionsanalysen ein Zusammenhang zwischen Entwicklungshilfeleistungen und Modernisierungserfolg gesucht. Je nach Auswahl der Indikatoren lassen sich positive als auch negative Effekte ausmachen (vgl. Stallings 1993: 1). Dies ist nicht weiter erstaunlich, da Entwicklungshilfe zum einen nur ein Faktor unter vielen ist, der Entwicklungsprozesse auslösen kann und zum anderen Entwicklungshilfe – gute wie schlechte – für die meisten Länder vom finanziellen Umfang her marginal ist. Und in vielen Ländern, in denen die finanzielle Unterstützung von außen einen hohen Anteil am Haushalt oder gar am gesamten BIP ausmacht, werden diese Hilfen zu-

3 Die GTZ will technische, wirtschaftliche und organisatorische Kenntnisse und Fertigkeiten vermitteln, um Menschen und Organisationen durch partnerschaftliche Zusammenarbeit in die Lage zu versetzen, ihre Lebensbedingungen zu verbessern. Auftragsvolumen ca. 1 Milliarde Euro/Jahr.

meist nicht investiv strukturbildend, sondern konsumptiv lebenserhaltend eingesetzt, so dass sie kaum entwicklungsfördernd wirken.

Die Befürworter der EZ, die Geberorganisationen, bevorzugen in der Regel andere Methoden, um den Erfolg ihrer Arbeit zu messen, nämlich Projekt- und Programmevaluationen. Dennoch überzeugen die Erfolgsberichte der Geberorganisationen nicht. Nun, das kann z.T. daran liegen, dass diese Broschüren nur als PR-Massnahmen gedacht sind und deshalb nicht die wahre Erkenntnis der EZ-Organisationen widerspiegeln. Aber es gibt tatsächlich ein Qualitätsproblem, nicht nur in der EZ, sondern auch in der Evaluation der EZ.

3. Defizite der Evaluation in der Entwicklungszusammenarbeit

Individualistische Fallbetrachtungen und international vergleichende Makrobetrachtungen scheinen wenig geeignet, die Wirksamkeit der EZ aufzudecken. Aber auch die bisher verwendeten Evaluationsinstrumente (die auf der Projekt- und Programmebene ansetzen) und ihre strategische Ausrichtung sind verbesserungsbedürftig:

- (1) Lange Zeit war es in der Entwicklungshilfe-Community üblich, den „Erfolg“ der Entwicklungspolitik am Input zu messen. Der Anteil der Entwicklungshilfenausgaben am Bruttoinlandsprodukt wurde als Orientierungsmaßstab verwendet. Noch immer ist zu beobachten, dass Länder, die das 0,7%-Ziel erreichen oder gar übertreffen, als vorbildliche Geber gelten. Wer diese Marke unterschreitet, muss sich den Vorwurf gefallen lassen, nicht genügend für die Lösung der Probleme der Dritten Welt zu leisten. Dennoch wird niemand bestreiten, dass Quantität nicht gleichzeitig Qualität bedeutet. Dieser Aspekt wird jedoch in der Diskussion um Geberquoten oft vernachlässigt.
- (2) Diese reine Inputbetrachtung wurde durch eine Input-Output-Betrachtung abgelöst, bei der die eingesetzten Mittel mit den erreichten Leistungen verglichen werden. Dabei werden z.B. die Ausgaben für Straßenbau mit der in Kilometern gemessenen Strecke fertiggestellter Straßen verglichen. Oder den Ausgaben für die Errichtung eines Krankenhauses wird die Zahl der zur Verfügung gestellten Betten gegenübergestellt. Doch solche Input-Output-Ziffern sagen nichts über die erzielten Wirkungen aus, also z.B. welche Folgen die Straße für die Umwelt hat, ob sich entlang der Straße Menschen angesiedelt haben, wer die Straße nutzt etc.; bzw. wer sich den Aufenthalt in den Betten des Krankenhauses leisten kann. Dies sind jedoch die entwicklungspolitisch entscheidenden Fragen.
- (3) In vielen Entwicklungsorganisationen hat bisher eine Planungs- und Implementationsperspektive vorgeherrscht, die sich stark an der Planerfüllung innerhalb der Förderlaufzeit orientiert. Hierfür wurde das geplante „Soll“ mit dem erreichten „Ist“ verglichen. Man konzentrierte sich auf Ziele und Ergebnisse, ohne ausreichend intendierte und nicht-intendierte Wirkungen zu erfassen. Deshalb kommt eine vom BMZ in Auftrag gegebene Studie zur Analyse der von deutschen EZ-Institutionen angewendeten Evaluationsverfahren zur Erfolgskontrolle zu dem Ergebnis:

„Wirkungs- und nachhaltigkeitsorientierte Erfolgskontrollen sind jedoch bei vielen EZ-Institutionen nicht oder nur in Ansätzen anzutreffen. Wirkungsmonitoring in der Verlaufsphase der Vorhaben ist oft nicht existent oder steht in den Anfängen. Die größten Defizite bestehen bei den Ex-post-Kontrollen“ (Borrmann u.a. 1999: 327).

Dass es sich hierbei nicht nur um ein deutsches Phänomen handelt, macht die Aussage von Bamberger (1991: 3) deutlich, der mit Blick auf die internationale Gebergemeinschaft feststellt: „(...) there is often very extensive monitoring of foreign aided projects during the period of implementation, but there is much less evaluation of how well projects operate, how effectively they are sustained, and to what extent they produce the intended impacts“.

- (4) Diese Aussagen machen auf ein weiteres Defizit aufmerksam, nämlich, dass sich bis heute noch immer viele Organisationen an der Erzielung kurzfristiger Wirkungen orientieren. Projekterfolg wird an dem gemessen, was im Rahmen der Förderlaufzeit erreicht wurde. Was nach der personellen und finanziellen Unterstützungsphase folgt, ob die Partner oder die Zielgruppen tatsächlich in die Lage versetzt wurden, ihre Probleme selbständig besser zu meistern als vorher, bleibt oft außerhalb des Blickfeldes. Da laufende Projekte eine künstliche Situation darstellen, in der externe Berater über hohe Ressourcen und eine große Interventionsmacht verfügen, entwickeln sich Projekte (bzw. die von ihnen angestoßenen Aktivitäten) nach dem Förderende nicht einfach linear weiter. Es entsteht eine neue Situation, in der sowohl ungeahnte Energien des Partners freigesetzt werden, als auch umgekehrt alle vom Projekt initiierten Aktivitäten zusammenbrechen können. Was tatsächlich geschieht, ist nur über ex-post Evaluationen herauszufinden. Doch diese werden nur sehr selten durchgeführt. Dies ist insoweit erstaunlich, als die Entwicklungspolitik vorgibt, an nachhaltigen Lösungen interessiert zu sein.
- (5) Während Qualitätsmanagement für in wettbewerbsoffenen Märkten agierende, gewinnorientierte Unternehmen eine hohe Priorität genießt, weil nur so Produkte und Dienstleistungen permanent verbessert werden können, ist die Akzeptanz von Qualitätsmanagement und Evaluation in vielen Geberorganisationen noch unterentwickelt. Deshalb werden oft nicht ausreichend Personalressourcen und Finanzmittel für die Durchführung von Evaluationen bereitgestellt. Statt dessen herrscht die Auffassung vor, das (gespendete) Geld müsse den Bedürftigen direkt zugute kommen. Viele sehen vor allem den Kostenfaktor, nicht aber das inhärente Lernpotenzial und die Chance, aufgrund von Evaluationsergebnissen Fehlentscheidungen zu vermeiden und damit Kosten zu sparen.
- (6) Die methodischen Möglichkeiten werden in EZ-Evaluationsstudien oft nicht genutzt. Die bisherige Evaluationspraxis – zumindest in Deutschland – erschöpft sich zumeist in Aktenanalysen und im Führen von Gesprächen mit Programmverantwortlichen auf verschiedenen Steuerungsebenen. Dabei werden häufig keine strukturierten Auswertungs- bzw. Interviewleitfäden verwendet oder gar standardisierte Interviews durchgeführt. Zielgruppenbefragungen werden selten vorgenommen, so dass die Perspektive der Betroffenen und potenziellen Nutznießer nur sehr vermittelt erfasst wird. In vielen Studien wird auf eine Darstellung der genutzten Methoden oder gar des Evaluationsdesigns verzichtet. Selten werden Begriffe definiert oder die Beurteilungsmaßstäbe offenge-

legt. Weiterhin wird in zahlreichen Studien die Kausalitätsproblematik kaum thematisiert. Ursachenzuschreibungen werden oft vorschnell und nicht immer nachvollziehbar vorgenommen. In der vom BMZ in Auftrag gegebenen Studie zur Analyse und Bewertung des Systems der Erfolgskontrolle in der deutschen EZ werden den Gutachtern „unzureichende oder gänzlich fehlende methodische Anstrengungen“ attestiert (Borrmann 1999: 71). Oft spiegeln die Evaluationen nicht belegte und deshalb intersubjektiv auch nicht nachvollziehbare persönliche Urteile der Gutachter wider, die mehr intuitiven Eingaben als empirisch überprüften Fakten folgen. Dabei dürften die stark subjektiv geprägten Urteile in hohem Maße von der Erfahrung, Professionszugehörigkeit und den Wertvorstellungen der Gutachter abhängen und schon deshalb von Fall zu Fall stark variieren.

- (7) Damit das gesamte Spektrum von Evaluationsinstrumenten angewendet werden kann, muss es den Gutachtern (den Evaluatoren) aber erst einmal bekannt sein. Doch dies kann nicht einfach vorausgesetzt werden, denn im deutschsprachigen Europa gibt es bisher keinen (postgraduierten) Studiengang, der zum Evaluator ausbildet und die notwendigen theoretischen, methodischen und praktischen Kenntnisse und Fertigkeiten vermittelt. Auch ein ausreichendes Weiterbildungsangebot ist nicht vorhanden und entwickelt sich erst in Ansätzen: An der Universität Bern wird im Wintersemester 2002/03 zum ersten Mal ein Ergänzungsstudium „Evaluation“ angeboten. In Deutschland erarbeitet der Arbeitskreis „Aus- und Weiterbildung“ der Deutschen Gesellschaft für Evaluation die Grundlagen für ein Studium der Evaluation. Bisher eignen sich Evaluatoren ihre evaluationsspezifischen Kenntnisse vor allem auf der Basis eines fachbezogenen Studiums an, indem sie durch einen „learning by doing“-Prozess Methoden der empirischen Sozialforschung oder der Ökonomie im Rahmen der spezifischen Kontextbedingungen von Evaluationen anwenden. Dies erschwert vor allem jungen Nachwuchskräften den Einstieg in den Evaluationsbereich.
- (8) In den Drittweltländern fehlen nicht nur Aus- und Weiterbildungsangebote im Bereich Evaluation, sondern auch erfahrene Gutachter. Deshalb ist die Ausbildung von Evaluatoren in Entwicklungsländern eine besonders wichtige Aufgabe, wenn Evaluationen partnerschaftlich durchgeführt werden sollen. Dies ist nicht nur deshalb notwendig, um zu stark geberzentrierte Evaluationen zu überwinden, sondern vor allem auch, weil die sozio-kulturelle Kompetenz und die Landeskenntnisse der einheimischen Evaluatoren unverzichtbar sind.
- (9) Evaluationen in der EZ sind zumeist interne Evaluationen oder werden im Auftrag von Geberinstitutionen durchgeführt. Was Bamberger (1991: 237) für die U.S.-amerikanische Evaluationsliteratur feststellt, gilt deshalb auch für die deutschsprachigen Länder: „Most of the literature relating to evaluation in developing countries is produced by, or for, donor agencies and is mainly concerned with the donor’s perspective“. Hinzu kommt, dass Geberorganisationen aufgrund ihrer Monopolstellung die Datenzugänge kontrollieren, die für empirische Studien notwendig sind. Zwar werden viele Studien und Gutachten unter Mitarbeit von Wissenschaftlern erstellt, doch da die Entwicklungshilfegeber viele Studien nicht freigeben, sind sie der wissenschaftlichen Diskussion weitgehend entzogen. Dies hat zum einen zur Folge, dass Ergebnisse nicht transparent gemacht werden, zum anderen führt die fehlende Kritik an den Evaluatio-

nen zu mangelnder Qualität. Während wissenschaftliche Studien einer professionellen Beurteilung standhalten müssen und Kritik zu einer Weiterentwicklung von Theorien und Methoden führt, kann sich ein solcher Prozess im Evaluationsbereich der EZ nur schwer entfalten.

- (10) Neben einer eingeschränkten Transparenz ist auch eine weithin fehlende Unabhängigkeit der Evaluationen festzustellen. Die meisten Evaluationen in der deutschen EZ werden nicht nur intern, sondern auch von den operativen Abteilungen selbst durchgeführt. Dies hat zwar Vorteile für die Geber-Institution (schnell, kostengünstig, gute Kenntnisse über Evaluationsgegenstand etc.), bringt aber auch erhebliche Nachteile mit sich (eingeschränkte Methodenkompetenz und Objektivität, fehlende Außensicht etc.).

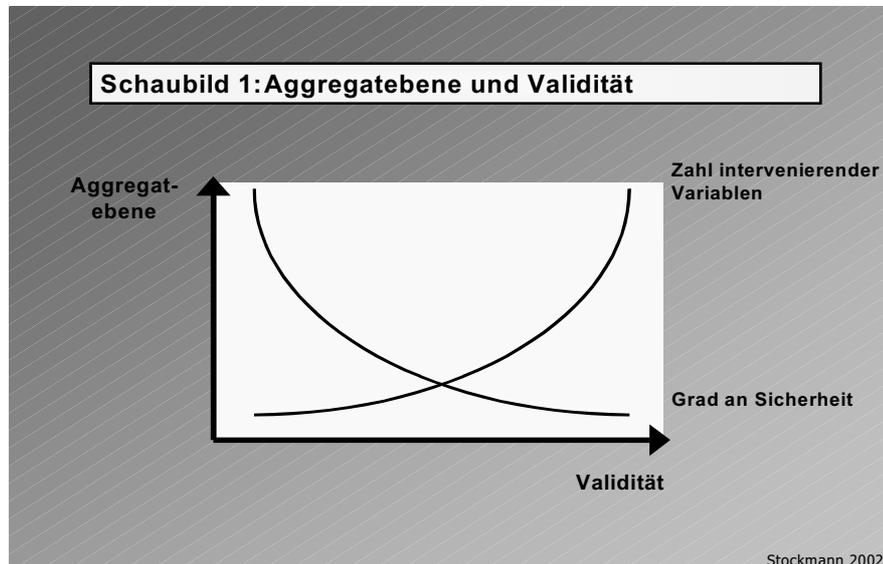
Interne Evaluationen sind zwar unabdingbar notwendig, aber ohne das Ferment von Fremd- oder externen Evaluationen werden nicht nur wichtige Lernpotenziale verschenkt, sondern ist auch die Glaubwürdigkeit der Evaluationsergebnisse eingeschränkt.

Es braucht nicht besonders erwähnt zu werden, dass dieser Defizitkatalog nicht für alle Organisationen der EZ gleichermaßen gilt. Die EZ ist charakterisiert durch eine äußerst große Vielfalt an staatlichen, halb-staatlichen und vor allem nicht-staatlichen Geberorganisationen, einer nicht minder heterogenen Vielfalt an Durchführungs- (Träger)organisationen in den Partnerländern und komplexen Programmstrukturen.

Dennoch scheinen die aufgezeigten Defizite weit verbreitet und mit fallweisen Abstrichen durchaus für die Evaluation in der EZ charakteristisch zu sein. Bestätigt wird diese Auffassung durch die hier bereits mehrfach zitierte, vom BMZ in Auftrag gegebene Studie zu den Verfahren der Erfolgskontrolle in der deutschen EZ (Borrmann u.a. 1999 u. 2001).

4. Objektive Evaluationsgrenzen und neue Herausforderungen

Selbst wenn es gelingt, die genannten Defizite in den nächsten Jahren allmählich zu überwinden, um die Evaluation in der EZ effektiver und wirksamer zu gestalten, werden immer auch objektive Grenzen die Evaluationsbemühungen einschränken. Wie eingangs erwähnt, besteht ein Kernproblem der Evaluation in der Erfassung von Wirkungen, der Analyse von Zusammenhängen und der Ursachenforschung. Dieses Problem lässt sich im direkten Projekt- oder Programmumfeld noch am ehesten eingrenzen und lösen. Mit zunehmender Aggregationsstufe, also höheren Handlungsebenen, steigt die Zahl der potenziell zu berücksichtigenden Einflussvariablen drastisch an. Damit nimmt aber die Wahrscheinlichkeit, mit Evaluationen zu gültigen Aussagen zu gelangen, deutlich ab. Die Ursache-Wirkungs-Ketten lassen sich immer schwieriger verfolgen.



Hinzu kommt, dass in der EZ, die klassischerweise für die Entdeckung von Ursache-Wirkungszusammenhängen angewendeten Methoden, nämlich Experimente, kaum eingesetzt werden können. Einerseits haben sich Experimente für die Analyse komplexer Phänomene als wenig geeignet gezeigt, zudem sind sie zu stark der internen Validität verhaftet und andererseits ist ihr Einsatz speziell in der EZ oft schon aus ethischen Gründen nicht möglich (vgl. Stockmann 1996: 107ff.). Allerdings stellen Experimente nicht die einzige Form dar, die Validität von Ergebnissen zu erhöhen. Die Verwendung theoretischer Konzepte und Hypothesen, mehrerer Datenerhebungsverfahren (Multimethodenansatz), die Bildung von Kontroll- oder zumindest Vergleichsgruppen und der Einsatz einer vergleichenden Analysekonzeption sind probate Mittel, die Gültigkeit von Aussagen zu erhöhen (vgl. Stockmann 1996: 117ff.).

Neuere gesellschaftliche und entwicklungspolitische Entwicklungen werden die Herausforderungen für die Evaluation zusätzlich verstärken. So wie die Ansprüche an die EZ steigen, steigen auch die Ansprüche an die Evaluation. Die EZ will sich nicht länger darauf beschränken, punktuelle, isolierte Verbesserungen bei einer Zielgruppe oder in einer Region zu erzielen, sondern strebt struktur- und systembildende Wirkungen an. Ob dies gelingt, kann aber zumeist nur auf einer höheren Aggregatebene (z.B. dem Bildungs- oder Beschäftigungssystem) festgestellt werden, was mit den genannten Evaluationsproblemen verbunden ist. Viele EZ-Interventionen zielen zudem auf dauerhafte Einstellungs- und Verhaltensänderungen (z.B. im Rahmen der Demokratieförderung, aber auch wenn es um die Bedeutung beruflicher Ausbildung, Hygiene oder Geburtenregulierung geht), die ebenfalls schwierig zu messen sind.

Des Weiteren haben sich auch die Standards, an denen der Erfolg der EZ gemessen werden soll, verändert. Es reicht nicht mehr aus, auf schlichte und leicht zählbare Input- und Outputziffern zu verweisen, sondern die Signifikanz, Effizienz,

Effektivität, Wirksamkeit und Nachhaltigkeit von Interventionen ist zu belegen. Nachhaltigkeit erfordert einerseits die Versöhnung ökonomischer, sozialer und ökologischer Ziele und andererseits eine Programmperspektive, die sich nicht nur damit zufrieden gibt, dass eine Zielgruppe oder ein Träger eine Neuerung in eigenem Interesse und zu ihrem/seinem eigenen Nutzen fortführt (vgl. Schaubild 2). Damit keine Insellösungen entstehen, muss eine Diffusion erfolgreicher entwicklungspolitischer Innovationen stattfinden, die möglicherweise sogar zu einer Leistungssteigerung gesamter Systeme beiträgt. Darüber hinaus müssen die Zielgruppen und Träger in die Lage versetzt werden, auf veränderte Umweltbedingungen flexibel und angemessen zu reagieren, sonst wird der einmal erreichte Entwicklungsstand zementiert und es findet keine weitere Entwicklung mehr statt.

Schaubild 2: Dimensionen der Nachhaltigkeit

Dimension	Typ	Merkmal
I	Projekt-orientiert	Zielgruppe/Träger führt Neuerungen in eigenem Interesse und zu eigenem Nutzen fort
II	output-/produktions-orientiert	Zielgruppe verfügt über eine Struktur, die sie/ihn in die Lage versetzt, Nutzen für andere dauerhaft sicherzustellen
III	system-orientiert	Die Innovation führt über Diffusionsprozesse zu einer Leistungssteigerung des gesamten Systems (z.B. des Bildungssystems oder des Gesundheitssystems)
IV	innovations-orientiert	Die Zielgruppe/der Träger verfügt über ein Innovationspotential, um auf veränderte Umweltbedingungen flexibel und angemessen zu reagieren

Stockmann 2002

Auch der Prozess der Globalisierung wird die Evaluation vor neue Herausforderungen stellen. Die weltweite Vernetzung ökonomischer Aktivitäten, die Verbreitung von Information, Wissen und westlichen Werten sowie die Globalisierung der Umweltprobleme setzen die Nationalstaaten unter starken Druck und schränken ihre Handlungsfähigkeit ein. Für die Evaluation bedeutet dies, dass es noch schwieriger wird, die Wirkungen von auf Nationalstaatsebene ansetzenden Programmen isolieren und Ursachenfaktoren zuschreiben zu können.

5. Perspektiven der Evaluation in der Entwicklungszusammenarbeit

5.1 Reformanstrengungen in Deutschland

Die Überwindung der genannten Defizite sowie die Anpassung an die Herausforderungen des neuen Jahrhunderts stellen die Zukunftsaufgaben zur Weiterentwicklung der Evaluation der EZ dar. Um die Nützlichkeit, Durchführbarkeit, Fairness und Genauigkeit – die vier zentralen Standards der Evaluation – zu verbessern (vgl. DeGEval 2002), ist eine umfassende Bestandsaufnahme notwendig, aus der wichtige Professionalisierungsschritte abgeleitet werden können.

In Deutschland hat das BMZ eine wichtige Rolle übernommen, indem es mehrere Maßnahmen zur Entwicklung einer Evaluationskultur bei deutschen Geber- bzw. Durchführungsorganisationen eingeleitet hat: (1) Das BMZ hat hierzu in einer unabhängigen Studie die von Institutionen der deutschen EZ praktizierten Erfolgskontrollen im Hinblick auf ihre Organisation, Konzeption, Methoden und Verfahren sowie Umfang und Struktur untersucht (Borrmann u.a. 1999). Dieser Bestandsaufnahme ist zwei Jahre später eine Evaluation der eingeleiteten Reformbemühungen gefolgt (Borrmann u.a. 2001). (2) Des Weiteren hat das BMZ für den staatlichen Bereich eine neue Evaluationskonzeption entwickelt, die auch GTZ und KfW mit einschließt. (3) Darüber hinaus hat das BMZ zum ersten Mal in seiner Geschichte eine breit angelegte Wirkungsanalyse ex-post durchgeführt, deren Ziel es war, „empirisch abgesicherte Aussagen zur Nachhaltigkeit der bundesdeutschen Entwicklungszusammenarbeit zu gewinnen“ (BMZ, Entwurf der QA vom Oktober 1999: 2). Hierzu wurden 32 Projekte in 19 Ländern aus vier Fördersektoren untersucht (vgl. BMZ 1999; Caspari, Kevenhörster und Stockmann 2000; Caspari und Stockmann 2001; Kevenhörster und Stockmann 2001; Stockmann 2001; Caspari, Kevenhörster und Stockmann i.E.).

Diese vom BMZ initiierten Aktivitäten fallen in einen Zeitraum, in dem sich die Evaluationsforschung in Deutschland zu professionalisieren sucht. Wie bereits in anderen europäischen Ländern hat sich auch in Deutschland (1997) eine Gesellschaft für Evaluation (DeGEval) als Fachverband für Evaluatoren sowie Auftraggeber und Nutzer von Evaluationen gegründet. Ein Arbeitskreis der DeGEval beschäftigt sich mit der „Evaluation von Entwicklungspolitik“ in dem sich mittlerweile fast 250 Personen eingeschrieben haben. Dieser Arbeitskreis stellt ein interdisziplinäres und interinstitutionelles Forum dar, das einen regelmäßigen Erfahrungsaustausch über aktuelle Entwicklungen und Probleme ermöglicht und eine Brücke zwischen Politik, Wissenschaft und Praxis schlagen will (vgl. Stockmann 2001).⁴

4 Adresse der AK Webseite: http://www.degeval.de/ak_ez/index.htm. Bisherige Themen der Tagungen waren:
– Evaluierbarkeit Entwicklungspolitischer Wirkungen.
– Aufgaben und Ziele des Arbeitskreises.
– Evaluation von Entwicklungspolitik.
– Evaluation: Reformmotor oder Reformbremse.

Zur Professionalisierung der Evaluation der EZ soll auch die Gründung eines wissenschaftlichen "Centrums für Evaluation" (2002)⁵ an der Universität des Saarlandes beitragen. Das Centrum, das eine Grundfinanzierung von der Saarländischen Landesregierung und der Universität des Saarlandes erhält und sich ansonsten über Drittmittel finanziert, konzentriert sich auf die Weiterentwicklung von Theorien und Methoden der Evaluation, bietet Qualifizierungsmaßnahmen an, berät öffentliche und private Stellen bei der Durchführung von Evaluationen und führt selbst drittmittelfinanzierte Evaluationen durch. Neben Umwelt- und Bildungsevaluation stellt die Evaluation der EZ einen inhaltlichen Schwerpunkt des Instituts dar. Außerdem wird an dem Centrum auch die erste interdisziplinäre, deutschsprachige „Zeitschrift für Evaluation“ herausgegeben.

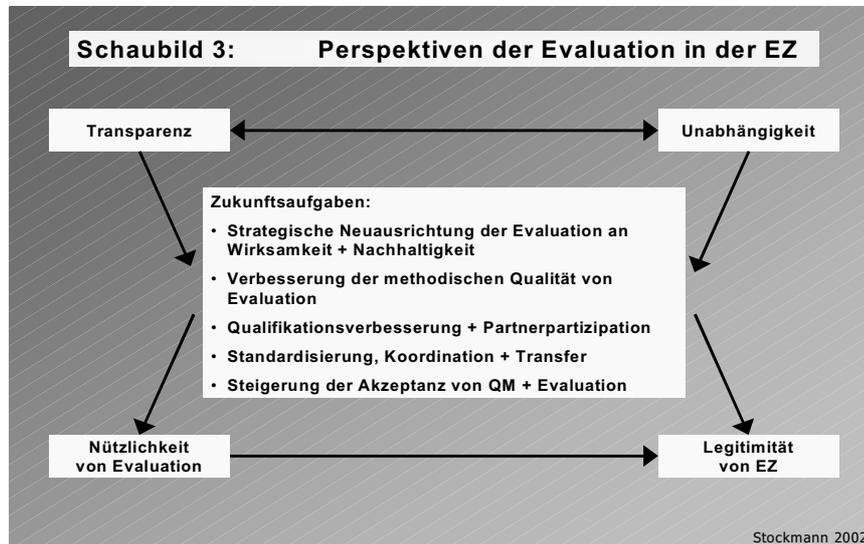
Schon seit einigen Jahren gibt es darüber hinaus in Deutschland die Forderung, ein politisch und administrativ unabhängiges Evaluationszentrum zu schaffen, das neben der praxisbezogenen Instrumentenentwicklung vor allem als Informationsbörse für Evaluationen in der EZ fungiert, indem es „best practices“ von Organisationen sammelt, dokumentiert auswertet und verbreitet, um das „institutional learning“ zu verbessern. Darüber hinaus wäre ein solches entwicklungspolitisches Zentrum dazu prädestiniert, unabhängig, objektiv und mit höchster Glaubwürdigkeit Berichte zur Wirksamkeit und Nachhaltigkeit der deutschen EZ zu erstellen. Dies würde dazu beitragen, die Legitimität der EZ insgesamt zu steigern. (Vgl. Stockmann 1996 u. 1998)

5.2 Professionalisierungsaufgaben

Eine Voraussetzung zur Professionalisierung der Evaluation in der EZ ist Transparenz. Diese ist notwendig, um die vorhandenen Defizite im Lichte der gestiegenen Ansprüche an die Evaluation der EZ erkennen zu können und um einen Dialog zu ermöglichen. Das Vorgehen des deutschen BMZ kann in dieser Hinsicht als vorbildlich bezeichnet werden, denn es hat durch eine schonungslose Offenlegung der in den entwicklungspolitischen Institutionen vorhandenen Defizite bei der Erfolgskontrolle die Basis für Veränderungen geschaffen. Der dadurch ausgelöste Reformdruck ist enorm und braucht Unterstützung, damit er erfolgreich in Veränderungsprozesse umgesetzt werden kann.

-
- Transparenz, Informationsfluss und follow up in Evaluierungsprozesse in der EZ.
 - Evaluationsbasiertes Lernen – Organisation von Lernen.
 - Qualitätsmanagement und Evaluation.

5 www.ceval.de



Was muss sich nach der hier vorgelegten Analyse im Einzelnen ändern?

(1) Strategische Neuausrichtung der Evaluation an Wirksamkeit und Nachhaltigkeit

Die einseitige Fixierung auf Planungs- und Implementationsprozesse sowie eine einseitige Soll-Ist-Beobachtung sollte einer an nachhaltiger Wirksamkeit ausgerichteten Evaluationsperspektive weichen. Hierfür ist das Wirkungsmonitoring für Projekte und Programme entscheidend zu intensivieren und die Zahl und die Qualität von Wirkungs- und ex-post durchgeführten Nachhaltigkeitsevaluationen ist deutlich zu erhöhen.

(2) Verbesserung der methodischen Qualität von Evaluation

Um die Qualität von Evaluation zu verbessern, sollten gewisse professionelle Mindeststandards erfüllt werden, die sich an den gängigen Standards der empirischen Sozialforschung sowie den DAC-Standards und den Standards der nationalen Evaluationsgesellschaften orientieren. Generell sollten anspruchsvollere Evaluationsdesigns als bisher angewendet werden, die nicht nur ein theoretisches Konzept oder zumindest einen konzeptionellen Rahmen erkennen lassen, die untersuchungsleitende Hypothesen zulassen, um kausale Ursachenzuschreibungen zu erleichtern. Außerdem sollte die Vielfalt empirischer Datenerhebungsverfahren ausgeschöpft werden. Zudem sind kulturell adaptierte sektor- und programmspezifische Evaluationsinstrumente und Methoden zu entwickeln.

(3) Qualifikationsverbesserung und Partnerpartizipation

Damit sich die Qualität der Evaluation verbessern kann, ist eine umfassende Aus- und Weiterbildung der in- wie ausländischen Evaluatoren notwendig. Nur wenn die Gutachter mit dem gesamten Spektrum von Evaluationsmethoden und ihren theoretischen Grundlagen vertraut sind, können sie beurteilen, welche Methoden bei welchen konkreten Aufgabenstellungen einzusetzen sind. Damit die Partnerländer in

Evaluationen aktiv eingebunden werden können, um vor allem auch eine geberzentrierte Sichtweise zu verhindern, ist die Aus- und Weiterbildung von Evaluatoren in den Partnerländern besonders voranzutreiben.

(4) Standardisierung, Koordination und Transfer

Um die Vergleichbarkeit von Evaluationsergebnissen zu erhöhen, um dadurch die Chancen für ein intra- und interinstitutionelles Lernen zu verbessern, sind gemeinsame Definitionen, Standards und Bewertungsmaßstäbe notwendig. Ein gemeinsames Evaluationsverständnis erleichtert zudem die gegenseitige Abstimmung bei Evaluationen sowie die Möglichkeiten, „joint evaluations“ durchzuführen. Dadurch steigt auch insgesamt die Chance, die bei Evaluationen gewonnenen Erkenntnisse (methodisch: best practices, inhaltlich: lessons learned) in den administrativen und politischen Entscheidungsgremien zu verbessern, um dadurch das „institutional learning“ zu erleichtern.

(5) Steigerung der Akzeptanz von Qualitätsmanagement und Evaluation

Qualitätsmanagement und Evaluation dürfen nicht länger nur als Kostenfaktor empfunden werden, den es möglichst niedrig zu halten gilt. Wer die Qualität und die nachhaltige Wirksamkeit von Projekten und Programmen verbessern will, der muss auch bereit sein, das dafür notwendige Qualitätsmanagement einzuführen und auf Wirkungen orientierte Evaluationen durchzuführen. Wenn Evaluationen nützlich sein sollen, müssen sie professionell durchgeführt werden und dürfen sich nicht ausschließlich an Kostenaspekten orientieren (QM und Evaluation gibt es nicht zum Spar- oder gar Nulltarif).

Transparenz ist notwendig, um Defizite erkennen und einen Dialog über Verbesserungsmöglichkeiten führen zu können. Dadurch wird es möglich, die Nützlichkeit von Evaluationen zu steigern. Nur dann, wenn Evaluationen professionell durchgeführt werden, entscheidungsrelevante Informationen liefern, rechtzeitig vorliegen und vertrauenswürdig sind, werden sie von Entscheidungsträgern als nützlich erkannt.

Nützliche und genutzte Evaluationen tragen auch automatisch zur Steigerung der Legitimität der EZ bei. Die Unabhängigkeit von Evaluationen allein kann die Legitimität von EZ nicht erhöhen. Dies haben die eingangs zitierten Betroffenheitsberichte deutlich gemacht. Ohne methodische Kompetenz sind auch unabhängige Studien unglaubwürdig und können nichts dazu beitragen, EZ effizienter, effektiver, wirksamer und nachhaltiger zu gestalten. Deshalb kann die Legitimität der EZ nur über die Steigerung der Professionalität der Evaluation und damit ihrer Nützlichkeit und Unabhängigkeit erreicht werden.

Da das Evaluationssystem der EZ bisher weitgehend über interne Evaluationen organisiert ist, muss nicht nur die Qualität der internen Evaluationen gesteigert werden, sondern sie muss auch um das Ferment qualitativ hochwertiger externer Evaluationen angereichert werden. Hier wird das Potenzial der sich mit Evaluation beschäftigenden Wissenschaft bisher kaum ausgeschöpft.

Die Nutzung vorhandener und die Erweiterung professioneller externer Evaluationskapazitäten würde wesentlich dazu beitragen, die Legitimität der EZ in der Politik und der Öffentlichkeit zu erhöhen. Um in Zeiten knapper Finanzmittel politische Ziele durchsetzen zu können, reichen Hinweise auf ethische oder moralische Verpflichtungen nicht aus. Wie in anderen Politikfeldern auch ist der empirisch

überprüfbarer Nachweis erforderlich, dass die angestrebten entwicklungspolitischen Ziele erreicht werden, dass die verausgabten Finanzmittel zu wirkungsvollen Effekten geführt und dazu beigetragen haben, nachhaltige Entwicklungsprozesse einzuleiten. Deshalb benötigt die Evaluation in der EZ mehr Transparenz, damit sich mehr Professionalität entwickeln kann und mehr Unabhängigkeit, damit das Vertrauen und die Glaubwürdigkeit in die Wirksamkeit und Nachhaltigkeit der EZ steigt.

6. Literatur

- Agarwal, Jamuna P.; Dippl, Martin; Glismann, Hans H. (1984): Wirkungen der Entwicklungshilfe (Forschungsberichte des BMZ Nr. 50) Köln.
- Bamberger, Michael (1991): The Politics of Evaluation in Developing Countries. In: Evaluation and Program Planning. Jg. 14: 325-339.
- Bauer, Peter T. (1984): Reality and Rhetoric. Studies in the Economics of Development. Cambridge, Massachusetts, London: Harvard University Press.
- BMZ (1999): Querschnittsauswertung: „Wirkungsuntersuchung abgeschlossener Vorhaben der deutschen EZ“. Entwurf. Gutachten im Auftrag des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), Bonn.
- Bode, Thilo (2002): Faire Chancen statt Almosen. In: Die Zeit Nr. 17, 18. April 2002. S. 24.
- Borrmann u.a. (1999): Erfolgskontrolle in der Deutschen Entwicklungszusammenarbeit: Analyse, Bewertung, Reformen Baden-Baden : Nomos-Verl.-Ges.
- Borrmann (2001): Reform der Erfolgskontrolle in der deutschen Entwicklungszusammenarbeit : eine Zwischenbilanz. Baden-Baden : Nomos-Verl.-Ges.
- Brüne, Stefan (Hg.), 1998: Erfolgskontrolle in der entwicklungspolitischen Zusammenarbeit. Hamburg: Deutsches Übersee-Institut.
- Cassen, Robert, 1990: Entwicklungszusammenarbeit. Bern, Stuttgart.
- Caspari, Alexandra; Kevenhörster, Paul; Stockmann, Reinhard (2000): Langfristige Wirkungen der staatlichen EZ. Ergebnisse einer Querschnittsevaluierung zur Nachhaltigkeit. In: Entwicklung und Zusammenarbeit. 41. Jg., H. 10.
- Caspari, Alexandra; Stockmann, Reinhard (2001): Nachhaltigkeit deutscher EZ-Projekte. Eine operationale Nachhaltigkeitsdefinition und ihre Anwendung. In: epd-Entwicklungspolitik. 14/2001.
- Caspari, Alexandra; Kevenhörster, Paul; Stockmann, Reinhard (Im Erscheinen): Das Schweigen des Parlaments: Die vergessene Frage der Nachhaltigkeit deutscher Entwicklungszusammenarbeit. In: Aus Parlament und Zeitgeschichte.
- DeGEval (2002): Standards für Evaluation. Köln: DeGEval.
- Elsenhans, Hartmut (Hg.) (1984b): Nord-Süd-Beziehungen. Stuttgart u.a.: Kohlhammer.
- Elsenhans, Hartmut (1984a): Abhängiger Kapitalismus oder bürokratische Entwicklungsgesellschaft. Frankfurt, New York: Campus.
- Erlar, Brigitte (1985): Tödliche Hilfe. Freiburg: Dreisman.
- Hancock, Graham (1989): Händler der Armut. Wohin verschwinden unsere Entwicklungsmilliarden? München: Droemer Knauer.
- Kevenhörster, Paul; Stockmann, Reinhard (2001): Wissenschaftlicher Rigorismus oder praxisorientierter Pragmatismus? Zum Verhältnis zwischen Entwicklungspolitik und Wissenschaft. In: Entwicklung und Zusammenarbeit. 42. Jg., H. 4.
- KfW (1999): Ergebnisse der Finanziellen Zusammenarbeit. Fünfter Auswertungsbericht über geförderte Vorhaben in Entwicklungsländern. Frankfurt a.M..
- Kromka, Franz; Kreul, Walter (1991): Unternehmen Entwicklungshilfe. Samariterdienst oder die Verwaltung des Elends? Zürich: Edition Interfrom.
- Myrdal, Gunnar (1984): International Inequality and Foreign Aid in Retrospect.

- Gerald M.; Seers, Dudley (Hg.) (1987): *Pioneers in Development*. Oxford University Press: Oxford. A World Bank Publication.
- OECD (1999): *DAC development Co-operation Report 1998*. Paris.
- OECD (1992): *DAC Principles for Effective Aid* Paris.
- OECD (1998): *Review of the DAC Principles for Evaluation of Development Assistance*. Paris.
- OECD: *Improving Evaluation Practices. Best Practice Guidelines for Evaluation and Background Paper*. <http://www.oecd.org/puma/>
- Rossi, Peter H.; Freeman, Howard E. (1999): *Evaluation. A Systematic Approach*. 6. Aufl., Thousand Oaks u.a.: Sage.
- Simon, Gabriela (1991): *Von Bürgern und Armen*. In: *Die Zeit*, vom 8. November 1991.
- Stallings, Barbara (1993): *The New International Context of Development*. In: *Items*. Jg. 47, H. 1; S. 1-11.
- Stockmann, Reinhard (1996a): *Die Wirksamkeit der Entwicklungshilfe. Eine Evaluation der Nachhaltigkeit von Programmen und Projekten der Berufsbildung*. Opladen: Westdeutscher Verlag.
- Stockmann, Reinhard (1996b): *Defizite in der Wirkungsbeobachtung. Ein unabhängiges Evaluationsinstitut könnte Abhilfe schaffen*. In: *Entwicklung und Zusammenarbeit*, H. 8.
- Stockmann, Reinhard; Caspari, Alexandra (1998): *Ex-Post Evaluation als Instrument des Qualitätsmanagements in der Entwicklungszusammenarbeit. Gutachten im Auftrag der Deutschen Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH, Eschborn*.
- Stockmann, Reinhard; Gaebe, Wolf (Hg.) (1993): *Hilft die Entwicklungshilfe langfristig? Bestandsaufnahme zur Nachhaltigkeit von Entwicklungsprojekten* Opladen: Westdeutscher Verlag.
- Stockmann, Reinhard (2001): *Der Arbeitskreis „Evaluation von Entwicklungspolitik“ in der DeGEval – Vernetzung von Politik, Praxis und Wissenschaft. Ein Beitrag zur Verbesserung der Evaluation von Entwicklungspolitik*. In: *Deutsche Gesellschaft für Evaluation e.V.: Evaluation. Reformmotor oder Reformbremse*. Köln.