

Neubert, Susanne

**Research Report**

## Wirkungsevaluation in der Entwicklungszusammenarbeit: Herausforderungen, Trends und gute Praxis

Analysen und Stellungnahmen, No. 3/2010

**Provided in Cooperation with:**

German Institute of Development and Sustainability (IDOS), Bonn

*Suggested Citation:* Neubert, Susanne (2010) : Wirkungsevaluation in der Entwicklungszusammenarbeit: Herausforderungen, Trends und gute Praxis, Analysen und Stellungnahmen, No. 3/2010, Deutsches Institut für Entwicklungspolitik (DIE), Bonn

This Version is available at:

<https://hdl.handle.net/10419/199925>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



## Wirkungsevaluation in der Entwicklungszusammenarbeit: Herausforderungen, Trends und gute Praxis

Wirkungsevaluationen erhalten heute endlich die Aufmerksamkeit in der entwicklungspolitischen Zusammenarbeit, die ihnen gebührt. Auch die Methodendiskussion ist in den letzten Jahren deutlich voran gekommen, und internationale Netzwerke zur Verbesserung von Methoden wurden eingerichtet.

Eine explizite Zuordnung von Entwicklungstrends zu Entwicklungsmaßnahmen oder externen Faktoren gehört heute zur guten Praxis, auch bei qualitativen Verfahren. Dennoch ist ein Trend hin zu quantitativ-experimentellen Verfahren unübersehbar. Dabei sollte man sich bewusst sein, dass mit Hilfe quantitativer Verfahren zwar konsistente Wirkungsschätzungen erbracht werden können; wie diese Wirkungen entstanden sind und wie nachhaltig sie sind, wird hiermit aber nicht erklärt. Auch können von den Ergebnissen keine direkten Politikempfehlungen abgeleitet werden. Zudem kosten experimentelle Verfahren ein Vielfaches qualitativer Verfahren, und sie sind zudem nicht oft ein-

setzbar. Die *eine* ideale Methodik gibt es daher nicht, sondern sie muss immer dem jeweiligen Entwicklungsvorhaben und -budget angepasst werden. Bei größeren Programmen empfiehlt sich manchmal ein Methodenmix. Ein solch umfassendes Vorgehen ist allerdings oft zu teuer. Evaluationsvorhaben, bei denen das gemeinsame Lernen im Zentrum steht und solche, die auf der politischen Ebene stattfinden, werden daher auch in Zukunft überwiegend qualitative Verfahren anwenden. Obwohl auch diese in den letzten Jahren weiterentwickelt wurden, gibt es hier ebenfalls Verbesserungsbedarf. So sollten qualitative Verfahren strukturierter durchgeführt werden und mit größeren Stichproben erfolgen als bisher üblich. Durch solche Verbesserungen würden allerdings dann auch qualitative Verfahren teurer werden. Qualitative Vorgehensweisen sollten zudem besser systematisiert und dokumentiert werden, so dass sie leichter zugänglich sind.

### Herausforderungen für Wirkungsevaluationen

Mit Hilfe von Wirkungsevaluationen wird überprüft, inwieweit entwicklungspolitische Maßnahmen zu Veränderungen in einer Gemeinde, einem Land oder einer Region geführt haben. In der Entwicklungszusammenarbeit (EZ) gehören Wirkungsanalysen schon lange zu den Evaluationsvorgaben des Entwicklungshilfesausschusses der *Organisation for Economic Cooperation and Development* (OECD), dem *Development Assistance Committee* (DAC). In der Praxis wird dies jedoch erst seit wenigen Jahren ernsthaft versucht.

Aufgrund des hohen Anspruchs an die EZ, weltweite Armut zu lindern, und des Anstiegs der Armut in vielen Ländern Afrikas steht die Entwicklungszusammenarbeit unter einem hohen Legitimationsdruck. Die Debatte über die Wirkungsmessung ist daher mit der Aid-Effectiveness-Debatte verschränkt. Aber auch die methodischen Anforderungen an Wirkungsanalysen wachsen, denn die Komplexität der Programme wird größer, die Interventions-ebenen und Finanzierungsmodalitäten verändern sich. Diese Veränderungen schlagen sich auch in der Evaluationsmethodik nieder. Eine Blaupause für die *eine* beste Methodik gibt es nicht. Es gibt aber methodische Ansätze, mit deren Hilfe die Ansprüche im Prinzip erfüllt werden können.

Während es in der Aid-Effectiveness-Debatte um die Prinzipien und Verfahren wirksamer EZ geht und auch um die

Frage der Kohärenz, d. h. des Zusammenwirkens verschiedener Politiken, bleiben diese Aspekte bei der eher technisch geprägten Methodendebatte unberücksichtigt.

Dies wäre aber für die richtige Interpretation der Ergebnisse von Wirkungsanalysen oft wichtig. Es ist nämlich sehr viel einfacher, eine bestehende positive Entwicklung in einem Land effektiv zu unterstützen, als etwa eine negative Entwicklung umzukehren. Auch ist es einfacher, höhere entwicklungspolitische Wirkungen zu erzielen, wenn andere Politiken kohärent dazu sind. In der öffentlichen Wahrnehmung gilt jedoch der „Bruttotrend“, d. h. die Entwicklung, wie sie de facto und aufgrund sämtlicher Einflüsse stattfindet, und nicht die Nettowirkung von isolierten Maßnahmen, die z. B. negative Trends bremsen, aber die Gesamtentwicklung nicht umdrehen. Um Programmen der EZ gerecht zu werden, d. h. ihre Wirkung nicht zu über-, aber auch nicht zu unterschätzen, ist es wichtig, die verschiedenen Einflüsse, die die Gesamtentwicklung ausmachen, analytisch zu trennen, also eine Zuordnung zwischen Ursache und Wirkung vorzunehmen. Methodisch gesehen ist dies die eigentliche Schwierigkeit bei Wirkungsanalysen.

Mit einem einfachen Vorher-nachher-Vergleich kann man keine Wirkungen von Programmen messen, denn Veränderungen können genauso gut durch externe Effekte entstanden sein. Die Wirkungszuschreibung muss deshalb explizit erfolgen. Bei quantitativ-rigorosen Ansätzen erfolgt dies über Kontrollgruppen, durch die ermittelt wird,

was geschehen wäre, wenn die Maßnahme nicht durchgeführt worden wäre. Bei qualitativen Verfahren erfolgt die Zuordnung dagegen durch die Plausibilisierung oder Erläuterung der Zusammenhänge, die mit Hilfe von Expertenwissen und durch Einschätzungen der Zielgruppen gewonnen werden.

Eine umfangreiche Überprüfung des deutschen Evaluierungssystems im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) im Jahr 2009 ergab, dass neben strukturellen Reformen des Evaluierungssystems auch Verbesserungen in der Methodik essenziell sind. Nur wenn auch die Methoden fundiert und nachvollziehbar sind, kann die Glaubwürdigkeit in der Öffentlichkeit steigen.

### Prinzipien qualitativer und partizipativer Ansätze

Qualitative Verfahren bemühen sich um Aushandlung. Sie reflektieren Prozesse, Interessen und Maßnahmen, um sie dann zu bewerten. Bei qualitativen und partizipativen Verfahren werden die Beteiligten einer Maßnahme als mündige, handlungsfähige Subjekte gesehen, die deren Wirkung deuten und ihr gegenüber rationale Strategien zur Verbesserung entwickeln. Das Evaluationsteam nimmt dabei oft eine Moderatorenrolle ein, und die Bewertung wird gemeinsam vorgenommen.

Der Umgang mit der Zuordnungslücke spielte bei solchen Wirkungsevaluationen noch in den 1990er Jahren keine vorrangige Rolle, es kam hauptsächlich darauf an, dass die Evaluierung nützlich war. Lange Zeit ging man davon aus, dass die Nettowirkungen nicht tatsächlich messbar sind und gab sich mit Wirkungseinschätzungen zufrieden. Der Schwerpunkt lag auf dem Verstehen des Gegenstands. Inzwischen hat sich diese Situation geändert. Auch bei qualitativen Verfahren spielt es eine wichtige Rolle, ob die Ursache-Wirkung-Zusammenhänge zuverlässig erkannt werden können. Bei partizipativen Verfahren wird davon ausgegangen, dass die Authentizität der Wirkungsbeschreibung dann am höchsten ist, wenn sie von Zielgruppen vorgenommen wird, die die Wirkungen selbst erleben. Evaluationen können dann Ausgangspunkt für die gemeinsame weitere Entwicklungsplanung sein, sie regen zum gemeinsamen Lernen an und erzeugen *Ownership*.

In qualitativen Verfahren werden die Daten durch Triangulation bis zum Sättigungspunkt validiert. Dabei werden Ergebnisse mit Hilfe anderer Quellen überprüft und es wird so lange geforscht, bis keine relevanten neuen Informationen mehr hinzukommen. Dieser Sättigungspunkt bleibt zwar oft Theorie, aber in manchen zielgruppenorientierten Befragungen kann er auch viel schneller erreicht werden, als man annehmen könnte. Bei partizipativen Ansätzen gilt demgegenüber das Prinzip, nur so viele Daten zu erheben, die wirklich notwendig sind, um die Zusammenhänge zu verstehen. Dieses als „optimale Ignoranz“ bezeichnete Prinzip muss in der Praxis mit dem Sättigungsprinzip ausbalanciert werden. Dass Kontrollgruppen bei diesen Designs in der Regel nicht gebildet werden, ist einer der größten Kritikpunkte der Verfechter quantitativer Verfahren. Prinzipiell ist die Befragung von Kontrollgruppen jedoch auch mit qualitativen Ansätzen möglich. Dies sollte daher in Zukunft durchaus manchmal erwogen werden.

### Stärken und Schwächen qualitativer Ansätze

Qualitative Verfahren umfassen verschiedene Designs und eignen sich zur Evaluierung auf allen Ebenen, d. h. von der Bevölkerungsbis hin zur Politikebene. Sie beruhen auf der Kombination verschiedener Instrumente, z. B. auf semi-strukturierten Interviews mit unterschiedlichen Akteuren, auf Gruppendiskussionen und Dokumentenanalysen. Die Stärke qualitativer Verfahren besteht in dem Erzielen eines umfassenden Verständnisses, aus dem Empfehlungen für das weitere Vorgehen im Programm entwickelt werden. Bei partizipativen Verfahren liegt der Schwerpunkt auf dem gemeinsamen Lernen. Dieses ist wichtig, auch damit Eigenverantwortung für die Verbesserung von Programmen erzeugt wird.

Die Schwäche qualitativer Verfahren liegt in der oft geringen Systematisierung. Bei mangelhafter Durchführung fehlt ein expliziter Umgang mit der Zuordnungslücke. Typischerweise wird beim qualitativen Vorgehen Fachkompetenz größer geschrieben als Methodenkompetenz. Eine weitere Schwäche liegt aufgrund der häufig geringen Aggregierbarkeit in der dann ebenfalls geringen Generalisierbarkeit der Ergebnisse. Seit etlichen Jahren existieren jedoch auch systematische qualitative Verfahren, die standardisiert werden können und die z. T. auch eine gewisse Aggregation zulassen (z. B. PCIA, MAPP, PIM). Sie beruhen zumeist auf Gruppendiskussionen und gehen explizit mit der Zuordnungslücke um. So werden z. B. mit Punktbewertungen die Ergebnisse im Rahmen relativer Aussagen quantifiziert. Dabei beziehen inzwischen einige Verfahren auch Schritte zum Auffinden unerwarteter und negativer Wirkungen ein.

Ein Manko der systematischen qualitativen Verfahren war bis jetzt jedoch ihre relativ schlechte Dokumentation. Die AG Wirkungsanalyse der Deutschen Gesellschaft für Evaluierung (DeGEval) hat dieses Manko 2009 ausgeglichen und eine „Landkarte der Verfahren der Wirkungsanalyse“ erstellt, in der nun alle bewährten Verfahren beschrieben werden.

### Prinzipien quantitativ-experimenteller Ansätze

Im experimentellen Design werden nach dem Zufallsprinzip ausgewählte Stichproben der Zielgruppen eines Programms, die als Versuchsgruppen „mit Maßnahmen“ bezeichnet werden, mit Kontrollgruppen „ohne Maßnahmen“ unter Anwendung ökonomischer Verfahren verglichen (Mit-ohne-Vergleich). Bei umfangreicheren Verfahren werden diese beiden Gruppen sowohl vor als auch nach Ablauf des Programms diesem Vergleich unterzogen. Die Differenz in der *Performance*, die sich zwischen den Gruppen feststellen lässt, zeigt dann die Wirkungen des Programms an (*double difference*). Die Bedeutung dieser Verfahren besteht darin, dass nur sie eine rigorose Schätzung von Wirkungseffekten erlauben.

In der Praxis stellt die Definition der Kontrollgruppe die größte Herausforderung dar. Denn die Kontrollgruppe muss mit der „Versuchsgruppe“ vergleichbar sein, aber ihr darf über die gesamte Laufzeit des Programms hinweg keine Maßnahme ähnlicher Art zugute kommen. Abgesehen davon, dass diese Forderung in einem Ressort, in dem es um Eigeninitiative im Rahmen des Nachfrageprinzips geht, spontan widersinnig erscheint, ist sie praktisch oft

auch nur sehr schwer zu erfüllen. Strategien, diese Schwierigkeit zu umgehen, liegen z. B. darin, Kontrollgruppen durch Zielgruppen zu ersetzen, die weniger lange an einer Maßnahme beteiligt sind. Auch kann die Stichprobengröße erhöht werden, um anschließend diejenigen auszusortieren, die sich an der Maßnahme innerhalb des Vergleichszeitraums beteiligen und damit zur Zielgruppe werden. Dies bedeutet jedoch, dass höhere Kosten entstehen. Heute werden manchmal auch *Matching*-Verfahren angewendet, bei denen Vergleichspaare, z. B. zwischen Gemeinden oder auch zwischen Individuen mit und ohne Maßnahmen, gebildet werden. Dieses Vorgehen lässt wiederum kleinere Stichproben zu. Eine weitere Möglichkeit, die Kosten zu senken, ist, Prototypen aus den Zielgruppen zu bilden und dann diejenigen auszuwählen, die diesen Prototypen entsprechen. Hierdurch kann mit wesentlich kleineren Stichproben gearbeitet werden; dies geht jedoch natürlich jeweils auf Kosten der wissenschaftlichen Rigorosität des Verfahrens.

Quantitative Wirkungsanalysen sind sehr voraussetzungsreich und sie verlangen mathematische Expertise, die zumeist nur an Lehrstühlen vorhanden ist. Auch verlangt das *double difference*-Verfahren die Planung des Designs bereits vor Beginn der Maßnahme. Der hauptsächliche Output bei solchen Verfahren besteht aus Tabellen, in denen für verschiedene Variablen die der Maßnahme zuzuschreibenden Veränderungen sowie deren Standardfehler aufgeführt sind. Die Zahlen zu interpretieren, verlangt erneute Expertise. Diese ist oft nicht vorhanden.

#### *Stärken und Schwächen quantitativ-experimenteller Ansätze*

Rigoreuse Verfahren können statistisch valide Wirkungsnachweise erbringen, und dies ist die Stärke des Ansatzes. Sie können aber nicht die Wirkungszusammenhänge erklären und daher auch keine Empfehlungen daraus ableiten, und dies stellt die wichtigste Schwäche dar. Hierzu sind qualitative Verfahren notwendig.

Experimentelle Evaluationen sind ungleich teurer als qualitative, v. a. weil sie mit repräsentativen Stichproben arbeiten und die Kontrollgruppe als Untersuchungseinheit hinzukommt. In der praktischen Erhebungsarbeit besteht der prinzipielle Unterschied zu partizipativen Verfahren darin, dass die Zielgruppen nicht selbst nach ihrer Einschätzung über die Wirksamkeit befragt werden, sondern dass lediglich die schieren Fakten, z. B. über das Budget eines Haushalts vor und nach der Maßnahme, abgefragt werden. Der Unterschied, der sich aus beiden Befragungen und zwischen Ziel- und Kontrollgruppen ergibt, zeigt dann die Wirkung der Maßnahmen an. Dies bedeutet, dass ein wichtiges Potenzial, die eigene Einschätzung der Partner und Zielgruppen über die Projektwirkungen, hier nicht genutzt wird. Stattdessen werden zahlreiche Daten erhoben, wobei sich erst bei der Auswertung zeigt, ob sie relevant sind. Hier besteht tendenziell die Gefahr der Schaffung von Datenfriedhöfen.

Experimentelle Verfahren eignen sich an sich nur für Projekte auf der Mikro- und z. T. auf der Mesoebene. Am leichtesten anzuwenden sind sie, wenn die angestrebten Wirkungen der Maßnahmen eindeutig und dichotom sind, wie es etwa in Impfprogrammen der Fall ist. Die Unterschiede der Versuchs- und Kontrollgruppen sind hier of-

fensichtlich, d. h. geimpft/nicht geimpft, erkrankt/nicht erkrankt. In der entwicklungspolitischen Praxis werden Zielsetzungen und erwartete Wirkungen von Maßnahmen jedoch zunehmend komplexer. Sollen daher solche Programme rigoros evaluiert werden, sind große Datensets erforderlich. Ein weiterer Nachteil experimenteller Ansätze besteht darin, dass i. d. R. Wirkungen erst am Ende der Maßnahme aufgezeigt werden. So kann die Evaluation nur Lerneffekte für zukünftige Programme erzielen und kommt dem evaluierten Programm selbst nicht zugute.

#### **Gute Praxis der Wirkungsevaluation**

Zunächst sollte eine Wirkungsevaluation nach allgemein gültigen Standards der Evaluation, z. B. wie sie die DeGEval aufgestellt hat, durchgeführt werden. Diese Standards heißen Vollständigkeit und Genauigkeit, Fairness, Transparenz und Nützlichkeit. Die ersten drei dieser Begriffe sprechen die mögliche Schiefelage bei Evaluierungen an, die durch Weglassung von Fakten und Gesichtspunkten, Ungenauigkeit der Erhebung oder Parteilichkeit von Evaluatoren entstehen kann. Nützlichkeit bedeutet, dass die Analyse Ergebnisse erbringen soll, die für die weitere Ausgestaltung des Programms verwertet werden können. Transparenz ist v. a. wichtig, um Bewertungen nachvollziehen zu können. Sowohl komplexe Tabellen, wie sie bei den quantitativen Evaluierungen oft entstehen, als auch selektive Plausibilisierungen, wie sie oft in qualitativen Verfahren vorgenommen werden, müssen daher transparent dargestellt und erläutert werden. Werden diese Prinzipien eingehalten, können alle seriösen Ansätze prinzipiell angemessen sein. Auf das Design und die Qualität der Durchführung kommt es an!

Bei der Methodenwahl sollte man sich am Programmziel, am Maßnahmentyp und am Zweck der Wirkungsanalyse orientieren. In der Regel sind die Programmziele der EZ komplex und nur mit mehreren Indikatoren beschreibbar. Auf internationaler Ebene stellen die *Millennium Development Goals* (MDGs) die wichtigsten Zielgrößen dar. Unterindikatoren sind anhand der Projektziele oder mit den Zielgruppen zu vereinbaren.

Bei qualitativen Wirkungsanalysen ist es wichtig, ergebnisoffen vorzugehen. Dies kann am besten gewährleistet werden, indem zuerst die Veränderungen vor Ort erhoben werden und erst dann die Zuordnung zu den Maßnahmen erfolgt. Lässt man sich dagegen zu sehr von den geplanten Wirkungen leiten, dann wird oft ein Tunnelblick eingenommen und das Programm nur in Bezug auf die gewünschten Wirkungen hin untersucht. Beide Blickwinkel einzunehmen – den kontext- und den maßnahmenbezogenen – und damit die Wirkungen letztlich einzukreisen, ist sicherlich die beste Vorgehensweise, aber sie ist natürlich auch aufwändig. Man muss sich bei Evaluationen in der Tat immer fragen, welche Datenfülle adäquat ist. Beide Gefahren sind real: Die der Produktion teurer Datenfriedhöfe und die der vorschleunigen Schlussfolgerungen aus unzureichenden Erhebungen.

Die Entscheidung über das genaue methodische Vorgehen ist auch vom zu evaluierenden Maßnahmentyp abhängig. Dabei ist noch wichtig zu berücksichtigen, ob die Wirkungen der direkten menschlichen Wahrnehmung unterliegen. Bei Maßnahmen, deren Wirkungen nur sehr ungenau



wahrgenommen werden können, wie z. B. zum Schutz der Biodiversität, sollten eher direkte Messungen angestellt werden. Zudem müssen für die Maßnahme klare Zielgruppen definiert werden können. Bei Maßnahmen der allgemeinen Wirtschaftsförderung sind dies z. B. mehrere Zielgruppen: Diejenigen, die direkt profitieren, wie z. B. Unternehmer und diejenigen, die letztlich profitieren, z. B. Konsumenten. Wirkungsanalysen von Politikberatung benötigen Analysen auf mehreren Ebenen. Hier muss zunächst die Wirksamkeit der Beratung, z. B. der Einflussgrad eines externen Beraters auf interne Entscheidungen und deren Umsetzung in einem Ministerium, und dann erst die Wirkung dieser Maßnahmen bei der Bevölkerung ermittelt werden. Das heißt, hier gibt es Wirkungen erster und zweiter Ordnung.

Letztlich entscheidet auch der Zweck der Evaluierung über die Methode. Soll das Lernen im Vordergrund stehen, sind qualitative Ansätze geeigneter. Bei großen Programmen, z. B. zur Erhöhung des Einkommens, kann die Integration quantitativ-experimenteller Ansätze sinnvoll sein.

Bereits vorhandene Informationen aus früheren Evaluationen, die Wirkungsaussagen enthalten, sollten in neue Analysen auf jeden Fall einbezogen werden. Dies wird leider oft versäumt, was Wissen und Ressourcen verschwendet. In Organisationen sollte das System der Wirkungsevaluation mit dem Monitoringsystem im Zusammenhang stehen. Es macht dabei Sinn, Prozessindikatoren von Wirkungsindikatoren zu unterscheiden.

Kommt ein Methodenmix zum Einsatz, müssen die Informationen aus den unterschiedlichen Quellen am Schluss zusammengeführt und verglichen werden. Hierzu eignen sich z. B. Ampel-, Benotungs- oder Punkteverfahren.

### Jüngste Entwicklungen

In den letzten Jahren haben sich durch eine komplexere EZ, die Paris-Erklärung zur Wirksamkeit der EZ, die DAC-Evaluierungskriterien und die regelmäßig durchgeführten Reviews die Ansprüche an Evaluationen erhöht. Ein großer Lern- und Verbesserungsbedarf ist in der Praxis bei qualitativen und quantitativen Evaluationen deutlich geworden.

Innerhalb der qualitativen Ansätze besteht ein großer Standardisierungs- und Systematisierungsbedarf. Für die Bevölkerungs- und Organisationsebene sind geeignete Verfahren vorhanden, sie müssen nur eingesetzt werden. Diese Verfahren sind z. T. sektoral und spezifisch, z. B. zur Analyse von Konfliktwirkungen, aber auch übersektoral anwendbar. Manche Verfahren beziehen zudem ungeplante Wirkungen in die Analyse ein. Andere arbeiten mit Matrices und Punktwerten, durch die Aussagen quantifiziert werden. Bei einigen Ansätzen wird die Zuordnungslücke heute explizit überbrückt (z. B. mit Hilfe einer Einflussmatrix). Hier kommt es darauf an, diese Verfahren endlich in der gängigen Praxis zu verankern und gut zugänglich zu machen. Entwicklungsbedarf besteht in der Methodik bei Wirkungsevaluationen auf der Politikebene, zur Politikberatung und bei der Evaluation neuer Finanzierungsmodalitäten (z. B. Budgethilfe). Die bestehenden Ansätze der Wirkungsevaluation hierfür sind noch lückenhaft, es existieren wenige dokumentierte Erfahrungen.

Um Verfahren weiter zu optimieren und zu vereinheitlichen, wurden 2006 und 2008 zwei internationale Netzwerke ins Leben gerufen: *Network of Networks on Impact Evaluation* (NONIE) und die *International Initiative for Impact Evaluation* (3ie). NONIE hat Geberorganisationen und multilaterale Banken als Mitglieder. 3ie ging aus einer Initiative des *Center for Global Development* (CGD) hervor. Wichtigstes Ziel der bisher eher auf quantitative Verfahren ausgerichteten Netzwerke besteht darin, vorbildhafte Wirkungsevaluierungen nach bestimmten Prinzipien durchzuführen und in einer Datenbank zu sammeln, auch um die als besonders wirksam erkannten Maßnahmen in künftige Projektplanungen einzubeziehen. Dadurch wird ermöglicht, die Erkenntnisse aus quantitativen Verfahren für die praktische EZ zu bündeln und daraus zu lernen.

### Methodenmix

Wirkungsanalysen können manchmal auch aus einer Kombination quantitativer und qualitativer Elemente bestehen, in der die jeweiligen Stärken der Ansätze kombiniert werden. Hierzu ist ein Team aus Methoden- und Sektorfachleuten erforderlich. Rigorose Untersuchungsdesigns sollten bewusst mit qualitativ-partizipativen Designs verschränkt werden. Während mit Hilfe des quantitativen Designs ein möglichst disaggregierter Wirkungsnachweis erbracht wird, z. B. im Hinblick auf soziale Gruppen, haben die qualitativen Erhebungselemente vielfältige Aufgaben. Zunächst können sie einen Überblick über die Rahmenbedingungen geben. Sie können außerdem die spezifischen Zusammenhänge und die Entwicklungsprozesse inklusive der möglichen Schwierigkeiten aufzeigen. Zudem können damit überraschende Ergebnisse aus den quantitativen Verfahren mit Hilfe der Aussagen von Zielgruppen erklärt werden und Empfehlungen für die Verbesserung des Entwicklungsprogramms abgeleitet werden. Ein solcher Methodenmix wird derzeit von der Deutschen Investitions- und Entwicklungsgesellschaft / Deutschen Gesellschaft für technische Zusammenarbeit (DEG / GTZ) zur Evaluierung des COMPACI-Programms (*Competitive African Cotton for Pro Poor-Growth*) der *Bill and Melinda Gates Foundation* in sieben afrikanischen Ländern entwickelt.



#### Dr. Susanne Neubert

ist Wissenschaftlerin am DIE (beurlaubt bis 2012) und derzeit für die Wirkungsevaluation des COMPACI in Sambia für die Deutsche Entwicklungsgesellschaft (DEG) in Zusammenarbeit mit dem National Opinion Research Center (Norc) an der University Chicago zuständig.

#### Literatur

Arbeitskreis „Evaluation von Entwicklungspolitik“ in der DeGEval – Gesellschaft für Evaluation (Hrsg.) (2009): Verfahren der Wirkungsanalyse – eine Landkarte für die entwicklungspolitische Praxis; online: [http://www.degeval.de/index.php?class=Calimero\\_Webpage&id=9037](http://www.degeval.de/index.php?class=Calimero_Webpage&id=9037)