

# RIGOROSE WIRKUNGSEVALUIERUNG IN DER DEUTSCHEN ENTWICKLUNGSZUSAMMENARBEIT

## Zusammenfassung

Experimentelle und quasiexperimentelle Wirkungsevaluierungen können einen wichtigen Beitrag für eine wirkungsvollere Entwicklungszusammenarbeit (EZ) leisten. In Deutschland wurden solche rigorosen Wirkungsevaluierungen (*rigorous impact evaluations*) bislang allerdings eher selten und unsystematisch durchgeführt. Zudem fehlt es an einer Strategie zur Nutzung daraus gewonnener Ergebnisse. Gleichzeitig sind Entwicklungen hin zu einer stärkeren Anwendung solcher Evaluierungen und verwandter Analysen erkennbar. Um das Potenzial rigoroser Wirkungsevaluierung systematisch in der deutschen EZ auszuschöpfen, sehen wir sechs Gestaltungsaufgaben: 1) Rigorose Wirkungsevaluierung und wirkungsorientierte Begleitforschung sollten ordnungspolitisch im deutschen EZ-System verankert und 2) dafür Ressourcen bereitgestellt sowie 3) die notwendigen Kapazitäten und Anreize in den EZ-Organisationen geschaffen werden. Darüber hinaus gilt es, 4) existierende Evidenz zu aggregieren, 5) die Wissenschaft in die Durchführung einzubeziehen und 6) auch in Partnerländern Kapazitäten zur Realisierung eigener rigoroser Wirkungsevaluierungen zu stärken.

## Wirkungsevaluierung in der deutschen EZ

Die Nachfrage nach fundierten Analysen zu Wirkungen entwicklungspolitischer Maßnahmen steigt an. Dies liegt unter anderem daran, dass entwicklungspolitische Themen an innergesellschaftlicher Relevanz gewonnen haben und in Zeiten von Migration, Gewaltkonflikten und Klimawandel Fragen nach nachweisbaren Wirkungen von humanitärer Hilfe und EZ noch wichtiger geworden sind.

Die Evaluierung der Entwicklungszusammenarbeit dient hierbei sowohl der Rechenschaftslegung gegenüber Parlament und Gesellschaft als auch Lernprozessen derjenigen, die für die politische Steuerung und die praktische Durchführung in der deutschen EZ verantwortlich sind. Evaluierungen tragen so zur evidenzbasierten Verbesserung von Strategien, Programmen und Projekten bei – und hierüber auch zur Legitimation des Politikfeldes.

Solch evidenzbasierte Politikgestaltung in der EZ stellt hohe methodische Anforderungen an Evaluierung und andere Formen angewandter Forschung. Entsprechend versierter sind die Analysen von Portfolios, Strukturen und Prozessen geworden. Dies gilt gleichfalls für die Wirkungsevaluierung auf der Ebene einzelner Projekte und Programme, wo rigorose (d. h. experimentelle und quasiexperimentelle) Methoden an Relevanz gewonnen haben. Die Bedeutung rigoroser Wirkungsevaluierung zur Armutsbekämpfung wurde zuletzt durch die Verleihung des Nobelpreises für Wirtschaftswissenschaften 2019 an Esther Duflo, Abhijit Banerjee und Michael Kremer, belegt, die maßgeblich zur Etablierung dieser Methoden in der Entwicklungsforschung beigetragen haben.

Rigorose Wirkungsevaluierung zielt auf eine kausale Zuordnung von möglichen Veränderungen durch eine entwicklungspolitische Maßnahme bei der Zielgruppe (z. B. Individuen, Haushalten, Schulen, Wirtschaftsakteuren) ab. Kern der Methode ist der Vergleich einer Interventionsgruppe, die eine bestimmte Intervention erhalten hat, und einer möglichst ähnlichen Kontroll- oder Vergleichsgruppe. Dadurch wird eine empirisch fundierte Aussage darüber ermöglicht, wie sich die Zielgruppe einer EZ-Intervention sowohl mit als auch ohne Durchführung der Maßnahme entwickelt hätte (Box 1).

### Box 1: Grundverständnis experimenteller und quasiexperimenteller Verfahren

Ähnlich wie bei experimentellen Verfahren in Medizin, Gesundheitsökonomie oder Erziehungswissenschaften basieren entwicklungspolitische Feldexperimente (auch *randomised controlled trials* oder RCTs) auf der zufallsbasierten Auswahl von Interventions- und Kontrollgruppe. Eine entwicklungspolitische Maßnahme wird bei einer Gruppe zufällig aus einer Grundgesamtheit gewählter Akteure durchgeführt (Interventionsgruppe). Die zufällig nicht gewählte Gruppe dient als Kontrollgruppe. Diese Vorgehensweise sorgt dafür, dass die beiden Gruppen in ihren beobachtbaren wie nicht beobachtbaren Eigenschaften als gleich erwartet werden können, sodass die Differenz der Zielgröße zwischen beiden Gruppen nach der Durchführung der Maßnahme einen unverzerrten Nettoeffekt schätzt. Quasiexperimentelle Verfahren können ähnlich verlässlich sein, indem sie mithilfe statistischer Verfahren Vergleichsgruppen bilden und Störgrößen systematisch ausschließen.

grundsätzlicher oder sektoraler Einschränkungen der Anwendbarkeit solcher Verfahren, unüberwindbarer ethischer Probleme oder nicht ausreichend theoriebasierter Vorgehensweisen als unbegründet, lösbar oder nur für eingeschränkte Bereiche als relevant erwiesen. Rigorose Wirkungsevaluierungen können in unterschiedlichen Sektoren theoriebasiert und ethisch vertretbar eingesetzt werden.

Gleichwohl sind rigorose Wirkungsevaluierungen nicht für alle wirkungsbezogenen Fragestellungen geeignet oder bei jeder Maßnahme das präferierte Verfahren. Gute anwendungsorientierte Forschung ist dadurch gekennzeichnet, dass die Methode dem praktischen Erkenntnisinteresse folgt und nicht umgekehrt. Auch bestehen erkenntnisstarke und nützliche Evaluierungen in aller Regel aus einer Kombination von Methoden; so können quantitative experimentelle Verfahren mit qualitativen Methoden kombiniert werden. Dabei sind rigorose Wirkungsevaluierungen besonders zielführend, wenn Fragen nach den Wirkungen einer Intervention im Fokus des Interesses stehen und eine entwicklungspolitische Maßnahme viele Akteure in ähnlicher Weise erreichen soll.

### Zur Verbreitung rigoroser Wirkungsevaluierungen

In den Sozialwissenschaften gehören diese Verfahren mittlerweile zum Standardrepertoire kausalitätsorientierter Feldforschung, und auch auf multilateraler Ebene sowie bei einigen bilateralen Geberorganisationen haben sie sich als ein relevanter Bestandteil der Evaluierungsportfolios etablieren können. Sie werden zur Rechenschaftslegung aber auch als wirkungsorientierte Begleitforschung durchgeführt, um zeitnah Informationen zur Verbesserung der Interventionen zu bekommen. Dabei können in mehrarmigen Wirkungsevaluierungen gleichzeitig mehrere Interventionen – auch kombiniert – getestet werden. Rigorose Evaluierungen werden ebenso in einigen Schwellenländern, zum Beispiel Indien, Mexiko oder Brasilien, vermehrt für die Untersuchung der Wirkungen staatlicher Sozialprogramme angewendet.

Darüber hinaus gibt es mittlerweile Verfahren, die die Erkenntnisse aus unterschiedlichen rigorosen Wirkungsevaluierungen mit ähnlichen Fragestellungen organisieren und aggregieren (Box 2). Insgesamt haben sich daher Bedenken hinsichtlich

In deutschen Entwicklungsorganisationen ist der Anteil rigoroser Wirkungsevaluierungen am Gesamtportfolio bislang allerdings vergleichsweise gering. Bisher wurden in der Kreditanstalt für Wiederaufbau (KfW) rigorose Wirkungsevaluierungen erprobt, die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) hat in circa 40 Maßnahmen (quasi-)experimentelle Verfahren verwendet und einzelne Nichtregierungsorganisationen haben erste rigorose Wirkungsevaluierungen durchgeführt.

Zudem ist bisher keine systematische Verwendung von Ergebnissen rigoroser Wirkungsevaluierung entlang einer übergeordneten Strategie des Lernens oder der Rechenschaftslegung zu erkennen. Vielmehr entsprang die Durchführung und Nutzung in vielen Fällen dem Engagement einzelner Akteure, die versucht haben, für einzelne Vorhaben einen Erkenntnisgewinn zu erzielen. Entsprechend bleibt ein solcher in den meisten Fällen auf das jeweilige Vorhaben und die Organisation beschränkt. Schließlich gibt es bislang keine Anreizsysteme, die die Anwendung rigoroser Verfahren und die Nutzung der daraus resultierenden Erkenntnisse systematisch fördern.

### Box 2: Die vier Phasen rigoroser Evidenz in der EZ

Die Bereitstellung rigoroser Evidenz in der EZ lässt sich international in vier Phasen darstellen:

- 1) Seit den 1990er Jahren führen Geber vermehrt ein Outcome-basiertes Ergebnismonitoring ein. Entlang eines solchen Results-Based-Managements werden Zielindikatoren formuliert und in deren Monitoring investiert.
- 2) Seit den 2000ern werden verstärkt rigorose Wirkungsevaluierungen durchgeführt. Diese sollen die Wirkungen von EZ-Interventionen zuverlässig identifizieren.
- 3) Um projektübergreifendes Lernen zu ermöglichen, werden seit den 2010er Jahren vermehrt systematische Reviews zur Aggregation der Ergebnisse aus einzelnen rigorosen Wirkungsevaluierungen genutzt. Durch die Aggregation wird die Validität von Aussagen erhöht. Zudem wird in sogenannten Evidenz-Karten (*evidence gap maps*) die existierende Evidenz in einzelnen Themenfeldern abgebildet.
- 4) In einer beginnenden vierte Phase wird die Nutzbarkeit existierender Erkenntnisse verbessert. Evidenz-Portale sollen das vorhandene Wissen der globalen Gemeinschaft einfach und zweckdienlich zur Verfügung stellen.

Vgl. hierzu White 2019

### Zukünftige Gestaltungsaufgaben

Im Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), in KfW und GIZ sowie in einigen Nichtregierungsorganisationen sind gegenwärtig Veränderungstendenzen hinsichtlich des Umgangs mit rigoroser Wirkungsevaluierung zu erkennen. Dies zeigt sich etwa in der Aufarbeitung existierender Erfahrungen mit diesen Methoden sowie einem verstärkten Interesse an rigoroser Wirkungsevaluierung, systematischen Reviews und Evidenz-Karten in ausgewählten Themenfeldern (z. B. Konfliktprävention und Übergangshilfe; Ausbildung und Beschäftigung).

Um das Potenzial rigoroser Wirkungsevaluierung für Lernprozesse in Entwicklungspolitik und Durchführungspraxis systematisch auszuschöpfen, sollten sechs strategische Gestaltungsaufgaben umgesetzt werden.

- Erstens bedarf es der ordnungspolitischen Verankerung rigoroser Wirkungsevaluierung und wirkungsorientierter Begleitforschung mit ihren Trägern und deren Kompetenzen im Gesamtsystem der deutschen EZ. Den entstehenden Leitlinien zur Evaluierung in der deutschen EZ kommt hierbei eine besondere Bedeutung zu.
- Zweitens bedarf es mehr finanzieller Ressourcen für die Durchführung und Nutzung rigoroser Wirkungsevaluierung in den staatlichen und nichtstaatlichen Entwicklungsorganisationen beispielsweise im Rahmen eines gezielten Förderprogramms.

### Gestaltungsaufgaben zur Nutzung der Potenziale von RIE



- Drittens ist es notwendig, in den Zentralen und Außenstrukturen der Entwicklungsorganisationen fachliche Kapazitäten und Anreizsysteme zu schaffen, die die Durchführung rigoroser Wirkungsevaluierungen begünstigen und sicherstellen, dass existierende Evidenz genutzt wird.
- Viertens sollte bei der Durchführung von rigorosen Wirkungsevaluierungen und Begleitforschung enger mit nationalen und internationalen Hochschulen und Forschungseinrichtungen – auch in Partnerländern – kooperiert werden. Wichtige Aufgaben für Evaluierung und Wissenschaft sind dabei eine gegenstandsangemessene Verknüpfung unterschiedlicher Methoden und eine Annäherung an die Bedarfe der Praxis.
- Fünftens sollte organisationsübergreifend eine Aufbereitung der bisherigen Erfahrungen geleistet werden. Dazu zählt es, Ergebnisse aus rigorosen Wirkungsevaluierungen in einem Repository zu erfassen und verstärkt in entsprechende Synthesen zu investieren.
- Sechstens gilt es bei der Durchführung von rigorosen Wirkungsevaluierungen und Begleitforschung, auch Akteure und Institutionen in den Partnerländern zu stärken. Diese sollten mittelfristig ihre eigenen Programme oder diejenigen der Geber eigenständig analysieren und bewerten können.

Ein BMZ-finanziertes Forschungsprojekt am DEval zu rigoroser Wirkungsevaluierung erarbeitet – über die oben genannten grundlegenden Gestaltungsaufgaben hinaus – konkrete Empfehlungen für eine systematischere und gegenstandsangemessene Durchführung und Nutzung rigoroser Wirkungsevaluierung in der deutschen EZ. Es fördert zudem Netzwerkbildung und bietet Beratung für deutsche EZ-Organisationen.

### Box 3: RIE am DEval

Am DEval sind insbesondere quasiexperimentelle Verfahren ein zunehmend wichtiger Bestandteil des Methodenportfolios. Sie werden nicht isoliert angewandt, sondern bilden in ausgewählten theoriebasierten Evaluierungen ein methodisches Element, das mit weiteren Ansätzen verknüpft wird. Evaluierungssynthesen und Evidenz-Karten finden ebenfalls immer mehr Eingang in das Portfolio des Instituts. Zudem versteht sich das DEval als Plattform für den Erfahrungsaustausch zu rigoroser Wirkungsevaluierung zwischen EZ-Organisationen, der Wissenschaft sowie im Themenfeld arbeitenden internationalen Organisationen.

## Literatur

**Funk, E., L. Groß, J. Leininger und A. v. Schiller (2019)**, „Erkenntnisse aus der wirkungsorientierten Begleitforschung: Potential und Grenzen der rigorosen Wirkungsanalyse von Governance-Programmen“, Discussion Paper 13/2019, Deutsches Institut für Entwicklungspolitik, Bonn.

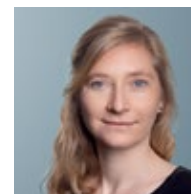
**White, H. (2019)**, „The twenty-first century experimenting society: the four waves of the evidence revolution“, Palgrave Communications 5, Nr. 47, Online verfügbar unter <https://www.nature.com/articles/s41599-019-0253-6>, zugegriffen 30.9.2019.



**Dr. Martin Bruder**  
Abteilungsleiter



**Prof. Dr. Jörg Faust**  
Direktor des Deval



**Dr. Marion Krämer**  
Teamleiterin

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