



NACHHALTIGKEIT IN DER DEUTSCHEN ENTWICKLUNGS- ZUSAMMENARBEIT

Evaluierungssynthese

Zusammenfassung

2018



DEval

DEUTSCHES
EVALUIERUNGSI
NSTITUT
DER ENTWICKLUNGS-
ZUSAMMENARBEIT

Impressum

Herausgeber

Deutsches Evaluierungsinstitut der
Entwicklungszusammenarbeit (DEval)
Fritz-Schäffer-Straße 26
53113 Bonn, Deutschland

Tel: +49 (0)228 33 69 07-0

E-Mail: info@DEval.org

www.DEval.org

Verfasst von

Dr. Martin Noltze

Dr. Michael Euler

Ida Verspohl

Verantwortlich

Prof. Dr. Jörg Faust (bis Juni 2016)

Dr. Sven Harten (ab Juni 2016)

Gestaltung

MedienMélange: Kommunikation!, Hamburg

www.medienmelange.de

Lektorat

Silvia Richter, mediamondi, Berlin

www.mediamondi.de

Bildnachweis

Gui Yongnian/123rf.com (Cover), FO Travel/Alamy Stock Foto
(Kap. 1), themacx/iStock.com (Kap. 2), Nikon'as/Fotolia.com
(Kap. 3), epicurean/iStock.com (Kap. 4 + 6), andresr/iStock.com
(Kap. 5), Steve Bloom Images/Alamy Stock Foto (Kap. 7)

Bibliografische Angabe

Noltze, M., M. Euler und I. Verspohl (2018),

*Evaluierungssynthese von Nachhaltigkeit in der deutschen
Entwicklungszusammenarbeit*, Deutsches Evaluierungsinstitut
der Entwicklungszusammenarbeit (DEval), Bonn.

Druck

Bonifatius,
Paderborn



© Deutsches Evaluierungsinstitut der
Entwicklungszusammenarbeit (DEval), Januar 2018

Druck ISBN: 978-3-96126-051-5

PDF ISBN: 978-3-96126-052-2

Das Deutsche Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) ist vom Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) mandatiert, Maßnahmen der deutschen Entwicklungszusammenarbeit unabhängig und nachvollziehbar zu analysieren und zu bewerten.

Mit seinen Evaluierungen trägt das Institut dazu bei, die Entscheidungsgrundlage für eine wirksame Gestaltung des Politikfeldes zu verbessern und die Transparenz zu den Ergebnissen zu erhöhen.

Der vorliegende Bericht ist auch auf der DEval-Website als pdf-Download verfügbar unter:
www.DEval.org/de/evaluierungsberichte.html

Anfragen nach einer gebundenen Ausgabe richten Sie bitte an:
info@DEval.org

ZUSAMMENFASSUNG

Hintergrund, Ziele und Evaluierungsgegenstand

Die Agenda 2030 für nachhaltige Entwicklung erhebt Nachhaltigkeit zum Leitbild globalen menschlichen Handelns. Die in der Agenda 2030 definierten nachhaltigen Entwicklungsziele (SDGs) verbinden wirtschaftlichen Fortschritt mit sozialer Gerechtigkeit und der schonenden Nutzung ökologischer Ressourcen. Dabei liegt die Umsetzung der Agenda 2030 in der Verantwortung aller Staaten. Gleichzeitig erfordert sie neue Kooperationen zwischen Politik, Privatwirtschaft, Wissenschaft und Zivilgesellschaft.

Auch die internationale Entwicklungszusammenarbeit (EZ) hat sich zu einer entsprechenden Neuausrichtung verpflichtet: Konzeption und Umsetzung von EZ-Maßnahmen müssen künftig den Zielen und Handlungsprinzipien der Agenda 2030 gerecht werden. Dies ist eine zentrale Herausforderung für die internationale EZ. Auf Ebene einzelner Maßnahmen bedarf dies insbesondere der Reflexion über soziale, wirtschaftliche und ökologische Wechselwirkungen sowie über Auswirkungen auf benachteiligte Gruppen. Um diesen Prozess zu unterstützen, sind evidenzbasierte Handlungsempfehlungen nötig. Derzeit gibt es nur eine begrenzte Anzahl an Vorhaben, die explizit in Anlehnung an die Agenda 2030 und deren Prinzipien konzipiert wurden. Eine empirische Auseinandersetzung mit der Nachhaltigkeit von EZ-Maßnahmen ist dennoch möglich.

Bereits seit 2006 wird die Nachhaltigkeit in Evaluierungen von Vorhaben der deutschen EZ systematisch überprüft und benotet. In dem Jahr verabschiedete das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) die „Evaluierungskriterien für die deutsche bilaterale Entwicklungszusammenarbeit. Eine Orientierung für Evaluierungen des BMZ und der Durchführungsorganisationen“. In Anlehnung an die vom Entwicklungsausschuss (DAC) der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) 1991 beschlossenen Prinzipien zur Evaluierung der Entwicklungszusammenarbeit enthält dieser Leitfaden Vorgaben zur Bewertung der Evaluierungskriterien Relevanz, Effektivität, Effizienz, übergeordnete entwicklungspolitische Wirkungen (Impact) und Nachhaltigkeit. Demnach wird die Nachhaltigkeit einzelner Maßnahmen anhand von verbindlichen Leitfragen bewertet. Als Ergebnis der Nachhaltigkeitsprüfung wird eine Note vergeben. Die Nachhaltigkeit von Vorhaben wird dabei

konzeptionell in enger Verbindung zur entwicklungspolitischen Wirksamkeit beurteilt. Daher kann erwartet werden, dass die bisherige Bewertungspraxis – über die entwicklungspolitische Wirksamkeit – bereits einige der Prinzipien der Agenda 2030 abdeckt.

Die hier durchgeführte Evaluierungssynthese zielt darauf ab, das Zusammenspiel verschiedener Determinanten bei der Nachhaltigkeitsbewertung von Vorhaben besser zu verstehen. Zweck der Untersuchung ist es, dazu beizutragen, die strategische und operative Ausrichtung der deutschen EZ besser an den neuen Anforderungen des modernen Nachhaltigkeitsverständnisses der Agenda 2030 auszurichten. Dies trägt der gesteigerten Bedeutung von Nachhaltigkeit im Kontext der SDG-konformen Evaluierung von Vorhaben in der deutschen EZ Rechnung.

Die vorliegende Evaluierungssynthese beinhaltet eine erste umfassende und systematische Auseinandersetzung mit der aggregierten Nachhaltigkeitsbewertung in Evaluierungen der deutschen finanziellen und technischen EZ. Dabei beschränkt sich die Untersuchung auf die Evaluierungs- und Bewertungspraxis der beiden großen staatlichen Durchführungsorganisationen (DO) – der KfW Entwicklungsbank (KfW) und der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ). Da der Evaluierungsgegenstand möglichst umfassend bearbeitet werden soll, wird er weder auf bestimmte Sektoren noch auf bestimmte Regionen oder Typen von Vorhaben beschränkt. Neben rein bilateralen Vorhaben in bestimmten Ländern sind auch Regional-, Sektor- und Globalvorhaben Teil der Untersuchung.

Methodisches Vorgehen

Die Analyse der Einflussfaktoren auf die vergebene Nachhaltigkeitsnote erfolgt durch multivariate Regressionsmodelle. Diese Modelle erlauben es, den Einfluss verschiedener Faktoren auf die zu erklärende Variable – die Nachhaltigkeitsnote von Vorhaben – zu bestimmen. Durch die begrenzte Datenverfügbarkeit können dabei nur bestimmte Einflussfaktoren berücksichtigt werden. Die Untersuchung beschränkt sich daher auf Merkmale der Vorhaben, Faktoren ihrer Implementierung sowie verfügbare Kontextinformationen. Zu Letzteren gehören sowohl Merkmale des unmittelbaren Kontextes der

Entwicklungsmaßnahmen als auch makroquantitative Indikatoren auf der Ebene der Partnerländer. Ferner greift die Analyse auf Ergebnisse der begleitenden Meta-Evaluierung zum Thema Nachhaltigkeit in der deutschen Entwicklungszusammenarbeit zurück. Die Ergebnisse der Meta-Evaluierung erlauben es einerseits, die zur Bewertung von Nachhaltigkeit herangezogenen Kriterien in den Analysen zu berücksichtigen. Andererseits wird die im Rahmen der Meta-Evaluierung durchgeführte Qualitätsbewertung der Evaluierungen als Gewichtungsfaktor für einzelne Beobachtungen in den Regressionsmodellen herangezogen. Dabei werden keine Beobachtungen von der Analyse ausgeschlossen, die Gewichtung einzelner Beobachtungen stellt jedoch sicher, dass die glaubwürdigsten Ergebnisse den größten Einfluss bei der Synthese erhalten.

Zentrale Ergebnisse, Schlussfolgerungen und Empfehlungen

Einflussfaktoren auf die Nachhaltigkeitsbewertung von Vorhaben

In der Evaluierungspraxis von KfW und GIZ variiert die vergebene Nachhaltigkeitsnote nur geringfügig. Über 84 Prozent der übermittelten Evaluierungen bewerten die Nachhaltigkeit mit der Notenstufe 2 oder 3. Dabei geht eine bessere Bewertung der DAC-Kriterien Relevanz, Effektivität, Effizienz und übergeordnete entwicklungspolitische Wirkungen (Impact) auch mit einer besseren Nachhaltigkeitsbewertung einher. So ist die Durchschnittsnote aller DAC-Kriterien (ohne Nachhaltigkeit) – nach statistischem Signifikanzniveau und nach Effektstärke – in allen Regressionsmodellen der wichtigste Einflussfaktor der Nachhaltigkeitsnote.

Nachhaltigkeit ist demnach ein übergeordnetes Evaluierungskriterium. Dieses enthält kaum genuine Einflussfaktoren, die trennscharf von den übrigen DAC-Kriterien abzugrenzen sind. Dennoch zeigen die Regressionsmodelle, dass bestimmte Faktoren bei der Notenvergabe besonders bedeutend sind. Insbesondere die aus der begleitenden thematischen Meta-Evaluierung gewonnenen Informationen erlauben Rückschlüsse auf die Nachhaltigkeit einzelner Maßnahmen. Die Ergebnisse der begleitenden Meta-Evaluierung zeigen außerdem, dass Nachhaltigkeit in der Evaluierungspraxis zwar anhand umfassender Kriterien bewertet wird, diese Bewertung aber gleichzeitig unsystematisch und uneinheitlich geschieht. Die Bewertung der Nachhaltigkeit steht zudem über die Bewertung der

entwicklungspolitischen Wirksamkeit auch immer im Zusammenhang mit der Bewertung der anderen DAC-Kriterien.

Auch die angewandten Evaluierungsformate bedingen Unterschiede in der Bewertung von Nachhaltigkeit. Während Ex-post-Evaluierungen ihre Bewertung auf Beobachtungen gründen, geschieht die Nachhaltigkeitsbewertung im Rahmen von Projektfortschrittskontrollen (PFK), Projektevaluierungen (PEV) und Schluss-Evaluierungen in Form einer Prognose. Im Vergleich zu den übrigen Evaluierungsformaten bewerten Ex-post-Evaluierungen die Nachhaltigkeit von Vorhaben tendenziell am schlechtesten. Doch nicht nur das Ergebnis der Bewertung, sondern auch die Kriterien, die der Bewertung zugrunde liegen, unterscheiden sich je nach Format. Ein Vergleich der Nachhaltigkeitsnoten zwischen verschiedenen Vorhaben ist somit nur eingeschränkt möglich. Generell kann aber festgehalten werden, dass in Ex-post-Evaluierungen die Rolle und die Beiträge der entwicklungspolitischen Partner und Zielgruppen von besonderer Bedeutung für die Nachhaltigkeit von Vorhaben sind. Demgegenüber werden bei der Nachhaltigkeitsbewertung in PFK, PEV und Schluss-Evaluierungen vor allem die direkten Leistungen und die Umsetzung einer Maßnahme sowie der Implementierungskontext berücksichtigt.

Neben diesen Unterschieden weisen die ermittelten Einflussfaktoren in den verschiedenen Evaluierungsformaten aber auch Gemeinsamkeiten auf. So wird der Absehbarkeit des Erhalts von Wirkungen sowohl in Ex-post-Evaluierungen als auch in PFK, PEV und Schluss-Evaluierungen ein signifikant positiver Einfluss auf die Nachhaltigkeit von Vorhaben zugesprochen. Dies zeigt, dass neben den Leistungen und Wirkungen von Vorhaben die Dauer von Wirkungen – ein konzeptionelles Kernelement der Nachhaltigkeitsbewertung – in allen Evaluierungstypen einen signifikanten Einfluss auf die Notenvergabe hat.

Empfehlungen im Sinne der Stärkung der Nachhaltigkeit von Vorhaben

Die im Folgenden genannten Empfehlungen ergeben sich aus den Ergebnissen und Schlussfolgerungen der Evaluierungssynthese. Aufgrund der Komplexität der Empfehlungen werden diese – in den jeweiligen Unterpunkten – durch Anregungen und Gedanken, die sich vornehmlich auf die Umsetzung beziehen, ergänzt.

Dem BMZ und den DO wird empfohlen, die Kapazitäten der Partner und Träger vor Ort bei der Planung und Durchführung von Vorhaben stärker zu berücksichtigen und systematisch zu fördern.

- In diesem Sinne könnte eine explizite Abschätzung der Kapazitäten aller relevanten Partner und Träger bereits bei der Planung von Vorhaben in das Votum zur Förderwürdigkeit eines Moduls einfließen. Dabei sollte sichergestellt werden, dass auf Seiten der Partner und Träger die technischen, finanziellen und institutionellen Voraussetzungen für die Fortführung der vormals durch die Maßnahme erbrachten Leistungen gegeben sind.
- Darüber hinaus könnte die Prüfung der Partner- und Trägerkapazitäten in regelmäßigen Abständen während eines laufenden Vorhabens wiederholt werden. Die Übertragung der Leistungen auf die Partner zum Ende eines Vorhabens könnte zudem durch das Entwickeln langfristiger Exit-Strategien sichergestellt werden.
- Durch die Stärkung des Partnersystems könnte die Eigenverantwortung der Partnerländer hinsichtlich der Umsetzung der Agenda 2030 sichergestellt werden.

GIZ und KfW wird empfohlen, steuerungsrelevante Faktoren eines Vorhabens zukünftig nicht nur im Hinblick auf die Wirksamkeit, sondern auch im direkten Bezug zur Nachhaltigkeit zu verstehen und zu berücksichtigen.

- Hierzu zählen insbesondere die Nutzung institutioneller Strukturen vor Ort, die Aufbereitung von Lernerfahrungen sowie das Entwickeln von Upscaling- und Exit-Strategien.

Systematisches Lernen aus Evaluierungen

Die Vergleichbarkeit von Evaluierungsergebnissen ist eine zentrale Voraussetzung für die Durchführung von Evaluierungssynthesen. Durch die Aggregation von Erkenntnissen aus einzelnen Evaluierungsberichten wird systematisches, strategisches und institutionenübergreifendes Lernen gefördert. Leider lassen sich die in den Evaluierungsberichten getätigten Aussagen zur Nachhaltigkeit von EZ-Maßnahmen nur bedingt vergleichen. Dies hat verschiedene Gründe:

Erstens bieten die Leitfragen zur Bewertung von Nachhaltigkeit zwar eine Orientierung bei der Notenvergabe, reichen für eine Operationalisierung aber nicht aus. Dies zeigt sich daran, dass die konkreten Bewertungskriterien, die sich hinter jeder einzelnen Note verbergen, vielfältig und nicht immer eindeutig zu benennen sind. Eine gewisse Flexibilität bei der Bewertung ist zwar aufgrund des diversen Portfolios an umgesetzten Maßnahmen notwendig; dennoch muss die Nachhaltigkeitsbewertung auch für Außenstehende nachvollziehbar und vergleichbar sein. Dieser Gedanke spiegelt sich mit dem Prinzip der gemeinsamen Rechenschaftslegung auch in der Agenda 2030 wider.

Zweitens weisen die hier betrachteten DO in der Bewertungspraxis und im Evaluierungsmanagement systematische Unterschiede auf. Die Ergebnisse belegen, dass Evaluierungen der GIZ – bei gleicher Anzahl positiv bewerteter Kriterien – signifikant bessere Nachhaltigkeitsnoten vergeben als Evaluierungen der KfW. Darüber hinaus führt die Anwendung verschiedener Evaluierungstypen sowohl innerhalb als auch zwischen den DO zu strukturellen Unterschieden in der Bewertung von Nachhaltigkeit. Auch im Evaluierungsmanagement der DO zeigen sich grundlegende Unterschiede. Bei der KfW werden alle Ex-post-Evaluierungen durch die Evaluierungsabteilung inhaltlich geprüft. Dabei wird die Bewertung einzelner Maßnahmen in den Kontext der Bewertung vergleichbarer Maßnahmen gesetzt. Demgegenüber liegt die Durchführung von PFK und PEV dezentral im Verantwortungsbereich des oder der Auftragsverantwortlichen einer Maßnahme. Während bei der KfW ein Kernteam an Mitarbeitenden alle Berichte kontrolliert und somit ein Mindestmaß an Vergleichbarkeit schafft, ist ein organisationsweiter Abgleich einzelner Berichte bei der GIZ im System dezentraler Evaluierungen nicht möglich. Es ist daher zu vermuten, dass Bewertungen von GIZ-Vorhaben insgesamt heterogener sind und stärker als bei der KfW von Eigenschaften der Autoren abhängen.

Drittens sind die von den DO aufbereiteten Meta-Daten von Evaluierungen und Vorhaben nur teilweise deckungsgleich. Für die vorliegende Analyse relevante Informationen wurden teilweise unvollständig oder aber nur von einer DO systematisch erfasst.

Die unzureichende Vergleichbarkeit der Nachhaltigkeitsbewertung erschwert es, nachhaltigkeitsförderliche Faktoren zu bestimmen. So ist beispielsweise anhand der vorliegenden Informationen nicht abschließend zu klären, ob die in den Modellen integrierten makro-ökonomischen und politischen Indikatoren tatsächlich keinen Einfluss auf die Nachhaltigkeit von Vorhaben ausüben oder ob aufgrund mangelnder Vergleichbarkeit und Transparenz der Bewertungsgrundlage kein Zusammenhang festgestellt werden kann. Das Potenzial zur Gewinnung strategischer und steuerungsrelevanter Erkenntnisse aus Evaluierungssynthesen ist somit stark eingeschränkt.

Empfehlungen hinsichtlich der Stärkung des systematischen, strategischen und institutionenübergreifenden Lernens

Auch die folgenden Empfehlungen werden durch Anregungen und Gedanken, die sich vornehmlich auf die Umsetzung beziehen, ergänzt.

Um die systematische Bewertung von Nachhaltigkeit zu gewährleisten, wird dem BMZ und den DO empfohlen, einheitliche und verbindliche Kriterien zu entwickeln. Diese sollten als Grundlage der Notenvergabe dienen und hierfür transparent gewichtet werden.

- Um dabei dem heterogenen Portfolio deutscher FZ und TZ gerecht zu werden, sollte auf angemessene sektor- oder regionalspezifische Flexibilität der Kriterien geachtet werden. Ein verbindlicher Umgang mit den Kriterien könnte gegebenenfalls auch sektoral oder für TZ-/FZ-Module getrennt festgelegt werden.

Dem BMZ und den DO wird empfohlen, Meta-Daten zu Vorhaben und deren Evaluierungen – soweit möglich – zwischen den DO zu harmonisieren und zentral zu erfassen.

- Eine systematische und zentrale Erfassung der Meta-Daten von Vorhaben und Evaluierungen würde institutionsübergreifende, aggregierte Analysen erheblich erleichtern und somit beschleunigen.
- Vor diesem Hintergrund könnten das BMZ und die DO prüfen, wie den Anforderungen der gemeinsamen Rechenschaftspflicht im Sinne der Agenda 2030 durch die Erfassung und Aufbereitung von Meta-Daten Rechnung getragen werden kann.