



Abschließende Mitteilung

an das
Bundesministerium für wirtschaftliche
Zusammenarbeit und Entwicklung

über die Prüfung

Evaluierung von Maßnahmen der
Entwicklungszusammenarbeit

Teil II – Maßnahmen der bilateralen Technischen Zusammenarbeit

Diese Prüfungsmitteilung enthält das vom Bundesrechnungshof abschließend im Sinne des § 96 Abs. 4 Satz 1 BHO festgestellte Prüfungsergebnis. Sie ist auf der Internetseite des Bundesrechnungshofes veröffentlicht (www.bundesrechnungshof.de).

Inhaltsverzeichnis

Abkürzungsverzeichnis	3
Begriffserläuterungen	4
0 Zusammenfassung	5
1 Vorbemerkungen	8
2 Ausgestaltung des Evaluierungssystems der GIZ	10
3 Vorbereitung der Evaluierung	12
3.1 Stichprobenverfahren	12
3.2 Finanzierung der Evaluierungen	15
4 Durchführung der Evaluierung	17
4.1 Ziel- und Ausgangswerte der Indikatoren	17
4.2 Evaluierungsansätze	19
4.3 Effizienzbewertung	21
4.4 Bewertung der Projekte	23
4.5 Qualitätsprüfung durch die Stabsstelle	25
5 Nutzen von Evaluierungen	27
5.1 Nutzen auf Projektebene	27
5.2 Nutzen oberhalb der Projektebene	29

Abkürzungsverzeichnis

AV	Auftragsverantwortlicher (Projektverantwortlicher im Partnerland)
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
DAC	Development Assistance Committee (Fachausschuss für Entwicklungszusammenarbeit der OECD (Organisation for Economic Cooperation and Development - Organisation für wirtschaftliche Zusammenarbeit und Entwicklung))
DeGEval	Gesellschaft für Evaluation e.V.
DEval	Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit gGmbH
DO	Durchführungsorganisation(en) für eine EZ-Maßnahme
EU	Europäische Union
EZ	Deutsche Entwicklungszusammenarbeit
FZ	Finanzielle Zusammenarbeit
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH
KfW	Kreditanstalt für Wiederaufbau
RIE	Rigorous Impact Evaluation (Methode für einen strengeren Wirkungsnachweis)
StbS	Stabstelle Evaluierung der GIZ
TZ	Technische Zusammenarbeit
Tz.	Textziffer
ZPE	Zentrale Projektevaluierung

Begriffserläuterungen

Evaluierung

Im Bereich der EZ sind Evaluierungen umfassende, systematische Analysen und Bewertungen von entwicklungspolitischen Projekten, Programmen, Instrumenten und Strategien. Sie dokumentieren und beurteilen, ob und wie Entwicklungsmaßnahmen ihre Ziele erreicht haben und ob sie als erfolgreich gelten können.

Evidenz

Evidenz in der EZ liegt vor, wenn die entwicklungspolitischen Wirkungen im jeweiligen Partnerland durch belastbare Daten und Erfahrungen belegt werden können. Die hierfür erforderlichen Daten müssen mit wissenschaftlichen Methoden erhoben und ausgewertet worden sein.

Kooperationspartner/Partnerländer

Kooperationspartner der EZ sind Partnerländer (Entwicklungs- und Transformationsländer) und ihre regionalen Zusammenschlüsse sowie die regionalen Wirtschaftskommissionen der Vereinten Nationen.

Länderstrategie

Länderstrategien bilden den strategischen Rahmen für EZ. Sie beschreiben die Zusammenarbeit mit dem Partnerland, die Aktivitäten anderer Geber im Land sowie die Schwerpunkte des deutschen Engagements. Sie definieren Ziele und Handlungsfelder für diese Schwerpunkte.

Monitoring

Monitoring ist eine systematische Überwachung von Prozessen, um festzustellen, ob diese den gewünschten Verlauf nehmen und ggf. steuernd eingreifen zu können.

Programme/Programmebene

Programme bündeln FZ- und/oder TZ-Projekte eines Sektors, die gemeinsame Wirkungen entfalten sollen. Programme sind operative Steuerungsinstrumente der deutschen Entwicklungszusammenarbeit (EZ).

Projekte/Projektebene

Projekte (das BMZ verwendet auch die Begriffe Module und Vorhaben) sind konkrete Maßnahmen der Finanziellen oder Technischen Zusammenarbeit (FZ/TZ).

0 Zusammenfassung

Mit seiner Prüfungsreihe „Evaluierung von Maßnahmen der deutschen Entwicklungszusammenarbeit“ untersucht der Bundesrechnungshof die Aufgaben und Rollen der einzelnen Akteure im Evaluierungssystem für die deutsche staatliche bilaterale Entwicklungszusammenarbeit (EZ). Im ersten Teil der Prüfungsreihe hat er geprüft, wie das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) die Evaluierungen plant und steuert und inwieweit es aus den Evaluierungen Nutzen für die strategische Ausrichtung der EZ zieht. Im vorliegenden zweiten Teil hat er die Evaluierung der bilateralen Technischen Zusammenarbeit durch die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH thematisiert. Die Prüfung beschäftigt sich nicht mit der Frage, ob die Evaluierungen den Anforderungen an eine Erfolgskontrolle im Sinne des § 7 Bundeshaushaltsordnung genügen.

- 0.1 Im Jahr 2017 führte die GIZ ein neues Projektevaluierungssystem – die Zentrale Projektevaluierung (ZPE) – ein. Danach werden Evaluierungen zentral von der Stabsstelle Evaluierung beauftragt und gesteuert und aus dem zu evaluierenden Projekt finanziert. Die Ausgaben der rund 100 ZPE pro Jahr betragen knapp 8 Mio. Euro (Tz. 2).
- 0.2 Nach dem Projektevaluierungssystem werden nicht alle Projekte evaluiert. Aus einer Grundgesamtheit der in einem Jahr zu beendenden Projekte, die das BMZ mit jeweils über 3 Mio. Euro finanziert, wird eine Stichprobe gezogen. Ziel der stichprobenweisen Evaluierung ist eine repräsentative Aussage über Erfolg oder Misserfolg der Grundgesamtheit (Erfolgsquote). Die Belastbarkeit der Stichproben für eine repräsentative Aussage soll Mitte 2021 überprüft werden. Nach Ansicht des Bundesrechnungshofes sollte künftig die Wirksamkeit der Projekte mit einer BMZ-Finanzierung auch unter 3 Mio. Euro ausreichend überprüft werden. Das BMZ teilt diese Empfehlung (Tz. 3.1).
- 0.3 Die Auftragsverantwortlichen der betroffenen Projekte vor Ort (AV) werden zurzeit mit einem Vorlauf von bis zu 18 Monaten über geplante Evaluierungen informiert. Grund hierfür sei, dass die AV die Evaluierung aus dem verbleibenden Projektbudget finanzieren und fachlich und organisatorisch vorbereiten müssten. Der Bundesrechnungshof hat empfohlen, die ZPE über eine Umlage auf alle Projekte zu finanzieren. Nach der Umstellung sollte die GIZ ihre AV ohne langen zeitlichen

Vorlauf über geplante Evaluierungen informieren. So kann die Sorgfalt bei der Projektdurchführung und dem kontinuierlichen Projektmonitoring hochgehalten werden. Für die aus dem Haushaltstitel der bilateralen TZ finanzierten Projekte hat das BMZ im September 2020 mit der GIZ eine Vereinbarung getroffen, wonach die ZPE durch Umlage über ein sog. Beitragsverfahren finanziert werden. Der Vorlauf zur Information der AV könne dadurch auf vier Wochen verkürzt werden (Tz. 3.2).

- 0.4 Nach den Vorgaben des BMZ soll die GIZ überprüfbare Ziele und angestrebte Wirkungen für ein Projekt definieren, deren Erreichen durch Indikatoren gemessen wird. Aus Sicht des Bundesrechnungshofes sind die Qualität der Ziel- und Ausgangswerte der im Projektvorschlag definierten Indikatoren entscheidend zur Beurteilung des Erfolges eines Projekts. Allerdings waren diese Werte der Indikatoren als auch die Wirkungsmodelle nicht immer hinreichend definiert. Die GIZ sollte bereits bei Erstellung von Projektvorschlägen hierfür strenge Maßstäbe anlegen. Das BMZ teilt diese Auffassung. Für die Prüfung neuer Projekte sei dies künftig Teil der Auftragsbedingungen. Zudem hat das BMZ inzwischen eine Verfahrensinformation zu Qualitätsstandards von Indikatoren in Kraft gesetzt (Tz. 4.1).
- 0.5 Nach den Vorgaben der GIZ soll die bei der ZPE angewandte Evaluierungsmethode ermöglichen, die Wirkungen eines Projekts zu erfassen. Es soll eine plausible Kausalbeziehung zwischen Maßnahme und Wirkung hergestellt oder zumindest der Beitrag einer Maßnahme zu Wirkungen belegt werden. Das BMZ sollte angesichts der Kosten jeder einzelnen ZPE darauf hinwirken, dass die GIZ bei Evaluierungen präzisere Aussagen über die Kausalbeziehung zwischen Maßnahme und Wirkung treffen kann. Auch das BMZ sieht im Verhältnis von Kosten und Aussagekraft der Evaluierung ein wichtiges Anliegen. Die GIZ werde sich weiter engagieren, den methodischen Ansatz zur Erfassung der Wirkungen zu verbessern (Tz. 4.2).
- 0.6 Die GIZ bewertet die Effizienz einer Entwicklungsmaßnahme u. a. durch Zuordnungen von Kosten zu Ergebnissen. Das dafür verwendete Excel-Tool scheint für eine nachträgliche Zuordnung geeignet. Der Bundesrechnungshof empfiehlt, dass die GIZ die Kosten und Ergebnisse einer Maßnahme bereits laufend IT-gestützt zuordnen sollte.

Nach Auskunft des BMZ arbeitet die GIZ daran, durch IT-gestützte Verknüpfung von Finanz- und Wirkungsmonitoring relevante Informationen zur wirtschaftlichen Steuerung der Projekte zur Verfügung zu stellen (Tz. 4.3).

- 0.7 Die Projekte werden nach international anerkannten Evaluierungskriterien bewertet. Für jedes Kriterium werden Einzelnoten vergeben. Diese fließen zu denselben Anteilen in die Gesamtbewertung ein. Hierdurch gleicht sich die Gesamtnote so aus, dass die weit überwiegende Zahl der Projekte „erfolgreich“ ist. Der Bundesrechnungshof hat eine unterschiedliche Gewichtung der Kriterien empfohlen. Dies ermöglicht eine differenziertere und aussagekräftigere Gesamtbewertung. Das BMZ hat inzwischen sogenannte „K.O.-Kriterien“ eingeführt. Wenn bestimmte Kriterien als „nicht erfolgreich“ bewertet werden, sei das Projekt insgesamt „nicht erfolgreich“ (Tz. 4.4).
- 0.8 Die GIZ überprüft die Qualität der Auftakt- und Evaluierungsberichte. Die Qualität einer Evaluierung bestimmt, ob die GIZ Zwischen- und Schlusszahlungen veranlasst. Die Genauigkeit wird mithilfe eines Prüfrasters nachgehalten. Prozessqualität und Nützlichkeit werden in Gesprächen thematisiert und über Checklisten erfasst. Der Bundesrechnungshof regt an, die Qualitätsstandards stärker formalisiert und früher zu überprüfen (Tz. 4.5).
- 0.9 Die GIZ gewinnt vermehrt und systematisch Erkenntnisse aus den Evaluierungen. Um den erstrebten Erkenntnisgewinn für künftige Entscheidungen zu erhöhen, sollte die GIZ ihre Evaluierungsergebnisse noch gezielter an alle Projektbeteiligten weitergeben und für Entscheidungen über Nachfolge-Projekte nutzen. Die in den ZPE ausgesprochenen Empfehlungen sollte die GIZ in einer Datenbank aufnehmen und thematisch auswerten. Sie sollte auch nachhalten, ob, inwieweit und in welcher Form die Adressaten den Empfehlungen nachkommen. Das BMZ teilt diese Empfehlungen. Bereits in der Auftaktphase einer Evaluierung werde künftig verstärkt auf Schlussfolgerungen für Folgeprojekte geachtet. Die GIZ möchte perspektivisch Evaluierungen IT-gestützt auswerten und Empfehlungen nachhalten. Darüber hinaus werde die GIZ ein geeignetes Verfahren der Nachverfolgung von Empfehlungen entwickeln und mit dem BMZ abstimmen (Tz. 5).

1 Vorbemerkungen

1.1 Prüfungsreihe „Evaluierung von Maßnahmen der deutschen Entwicklungszusammenarbeit“

Mit unserer Prüfungsreihe „Evaluierung von Maßnahmen der deutschen Entwicklungszusammenarbeit“ untersuchen wir das deutsche Evaluierungssystem für die staatliche bilaterale Entwicklungszusammenarbeit (EZ) und die Aufgaben und Rollen der einzelnen Akteure. Mit der Prüfungsreihe richten wir den Fokus vor allem auf eine systematische und das gesamte Förderspektrum umfassende Evaluierungstätigkeit. Neben dem wirksamen Einsatz von Haushaltsmitteln ist auch von Bedeutung, ob die Evaluierungen Erkenntnisse zur Nachhaltigkeit der EZ liefern. Die Prüfungsreihe beschäftigt sich nicht mit der Frage, ob die Evaluierungen den Anforderungen an eine Erfolgskontrolle im Sinne des § 7 Bundeshaushaltsordnung genügen.

Im ersten Teil der Prüfungsreihe haben wir zunächst Ihre Rolle näher betrachtet. Dabei sind wir insbesondere der Frage nachgegangen, in welcher Art und Weise Sie Evaluierungen für die Steuerung der EZ nutzen.¹ Im vorliegenden zweiten Teil der Prüfungsreihe haben wir die Evaluierung der bilateralen Technischen Zusammenarbeit (TZ)² durch die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH geprüft. In einem dritten Teil prüfen wir die Evaluierungen der bilateralen Finanziellen Zusammenarbeit (FZ)³ durch die Entwicklungsbank der Kreditanstalt für Wiederaufbau (KfW). Zudem ist als ein vierter Teil eine Prüfung der Evaluierungstätigkeiten des Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) geplant.

Für einen umfassenden Überblick über die Evaluierungstätigkeit der GIZ haben wir in dieser Prüfung Abläufe und Zuständigkeiten bei Evaluierungen und

¹ Vgl. unsere Abschließende Prüfungsmitteilung vom 27.04.2021, Geschäftszeichen II 3 – 2019 – 0210.

² Die TZ soll die Fähigkeiten von Menschen, Organisationen und Gesellschaften in den Partnerländern erhöhen und sie in die Lage versetzen, eigene Ziele durch effektiven, effizienten und nachhaltigen Einsatz von Ressourcen zu verwirklichen. Aufgabe der TZ ist auch der Aus- und Aufbau von Trägerstrukturen sowie die Unterstützung von Planungs-, Durchführungs- und Steuerungsinstitutionen bei Programmen der Kooperationspartner.

³ Durch die FZ sollen Investitionen der Kooperationspartner (d. h. Partnerländer und ihre regionalen Zusammenschlüsse sowie die regionalen Wirtschaftskommissionen der Vereinten Nationen) durch Finanzmittel und ergänzende Maßnahmen gefördert werden. Es sollen u. a. die soziale und wirtschaftliche Infrastruktur sowie der Umwelt- und Ressourcenschutz gefördert werden.

Monitoring⁴ während der Projektdurchführung untersucht.⁵ Vertieft haben wir neun Projektevaluierungen geprüft.⁶

Zu unserer Prüfungsmitteilung vom 31. August 2020 haben Sie und die GIZ mit Schreiben vom 14. Dezember 2020 und 27. Januar 2021 Stellung genommen. Unter Berücksichtigung Ihrer Stellungnahmen schließen wir diesen Teil der Prüfung mit den folgenden Feststellungen ab.

1.2 Begriff der staatlichen bilateralen Entwicklungszusammenarbeit

Die staatliche bilaterale EZ mit Partnerländern⁷ wird vom deutschen Staat initiiert, finanziert und organisiert.⁸ Sie umfasst alle Projekte⁹ der FZ und der TZ, die in völkerrechtlichen Verträgen direkt mit den Regierungen der Partnerländer vereinbart wurden.¹⁰ Für die von der deutschen Seite zu erfüllenden Aufgaben der Vorbereitung und Umsetzung der bilateralen entwicklungspolitischen Projekte bedient sich die Bundesregierung grundsätzlich deutscher Durchführungsorganisationen (DO).¹¹

⁴ Monitoring ist eine systematische Überwachung von Prozessen, um festzustellen, ob diese den gewünschten Verlauf nehmen und ggf. steuernd eingreifen zu können.

⁵ Neben Gesprächen mit der Stabsstelle Evaluierung und dem Fach- und Methodenbereich der GIZ haben wir Informationen aus dem Dokumentenmanagementsystem der GIZ gewonnen.

⁶ 35 ZPE waren zum Zeitpunkt unserer Erhebungen bereits abgeschlossen. Die Auswahl der näher untersuchten ZPE, haben wir zufällig getroffen, sie ist nicht repräsentativ.

⁷ Kooperationspartner der EZ sind Partnerländer (Entwicklungs- und Transformationsländer) und ihre regionalen Zusammenschlüsse sowie die regionalen Wirtschaftskommissionen der Vereinten Nationen; Quelle: „Leitlinien für die bilaterale Finanzielle und Technische Zusammenarbeit mit Kooperationspartnern der deutschen Entwicklungszusammenarbeit“, Verwaltungsvorschrift der Bundesregierung, S.5).

⁸ Projekte der bilateralen Technischen Zusammenarbeit werden aus Kapitel 23 01 Titel 896 03 und Maßnahmen der bilateralen Finanziellen Zusammenarbeit aus Kapitel 2301 Titel 866 11 und 896 11 finanziert. Daneben werden entwicklungswichtige Maßnahmen von nichtstaatlichen Organisationen von Ihnen (anteilig) finanziert.

⁹ Projekte (Sie verwenden auch die Begriffe Module und Vorhaben) sind konkrete Maßnahmen der Finanziellen oder Technischen Zusammenarbeit (FZ/TZ); Quelle: „Verfahrensablauf für die Planung und Durchführung von Maßnahmen der FZ und TZ“, Handreichung Nummer 023.

¹⁰ Die Bundesregierung sagt dem Kooperationspartner entweder bei Regierungsverhandlungen oder durch offizielle Mitteilung die Förderung geeigneter Entwicklungsmaßnahmen jeweils unter dem Vorbehalt des Abschlusses einer völkerrechtlichen Übereinkunft – auf der Grundlage von Rahmenabkommen – und der positiven Prüfung durch die Durchführungsorganisationen zu.

¹¹ DO in der FZ sind die KfW und die Deutsche Investitions- und Entwicklungsgesellschaft mbH (DEG). In der TZ ist dies die GIZ, soweit die Maßnahmen nicht von der Bundesregierung und ihren Dienststellen durchgeführt werden, (insbesondere Bundesanstalt für Geowissenschaften und Rohstoffe und Physikalisch Technische Bundesanstalt). Die DO präzisieren mit den zuständigen Stellen des Kooperationspartners die Einzelheiten der in den völkerrechtlichen Übereinkünften geregelten Förderungen, in der Regel in privatrechtlichen Verträgen.

1.3 Begriff und Zweck einer Evaluierung

Sie beschreiben den Begriff und den Zweck einer **Evaluierung** wie folgt:

„Im Rahmen der Entwicklungszusammenarbeit sind Evaluierungen umfassende und systematische Erfolgskontrollen von entwicklungspolitischen Programmen, Instrumenten und Strategien. Sie dokumentieren und beurteilen,

- *ob und wie Entwicklungsmaßnahmen ihre **Ziele** erreicht haben und*
- *ob sie als **erfolgreich** gelten können.“¹²*

Die Wirksamkeit der EZ ist nach sechs Schlüsselkriterien (**Relevanz, Effektivität, Effizienz, Kohärenz, übergeordnete entwicklungspolitische Wirkungen und Nachhaltigkeit**) zu bewerten. Auf diese Kriterien hat sich die internationale Gebergemeinschaft im Development Assistance Committee (DAC) geeinigt.¹³ Das DAC hat „Kohärenz“ erst im Dezember 2019 als neues Kriterium angenommen, daher wurde es in den unseren Erhebungen unterliegenden Evaluierungen der GIZ noch nicht eigenständig berücksichtigt.¹⁴

2 Ausgestaltung des Evaluierungssystems der GIZ

Zur Verbesserung der Planung und Steuerung und für eine verbindliche Festlegung und Dokumentation der Abläufe der EZ beschlossen Sie im Jahr 2017 nach Abstimmung mit den DO eine Gemeinsame Verfahrensreform.

In diesem Zusammenhang führte die GIZ die Zentrale Projektevaluierung (ZPE) ein. Nach diesem neuen System werden die Evaluierungen seit dem Jahr 2018 nicht mehr dezentral von den Projekt-/Auftragsverantwortlichen vor Ort (AV) in den Partnerländern beauftragt und gesteuert, sondern zentral von der Stabsstelle Evaluierung (StbS)¹⁵. Auch intern haben sich die Zuständigkeiten

¹² Vgl. <https://www.bmz.de/de/service/lexikon#lexicon=14352>; letzter Abruf: 6. April 2021.

¹³ Die sechs Schlüsselkriterien des DAC der Organisation for Economic Cooperation and Development (OECD) sind:

- o **Relevanz** (Tun sie das Richtige?),
- o **Effektivität** (Erreicht die Maßnahme das angestrebte Projektziel?),
- o **Effizienz** (Werden die Wirkungen wirtschaftlich erreicht? Tun sie es richtig?),
- o **Kohärenz** (Ist die Maßnahme mit anderen kompatibel?);
- o **Übergeordnete entwicklungspolitische Wirkungen** (Trägt die Maßnahme zur Erreichung übergeordneter Ziele bei?) und
- o **Nachhaltigkeit** (Sind die Wirkungen von Dauer?).

¹⁴ Die GIZ wird das Kriterium in ihren Evaluierungen ab Oktober 2020 verwenden.

¹⁵ Die frühere „Stabsstelle Monitoring und Evaluierung“ der GIZ heißt seit der Reorganisation 2016 „Stabsstelle Evaluierung“. Sie untersteht organisatorisch direkt dem Vorstand und ist vom operativen Geschäft getrennt.

geändert: Kernelemente der Reform sind die Trennung von projektvorbereitender Prüfung und projektbewertender Evaluierung. Die StbS agiert unabhängig von den operativen Einheiten; sie plant und steuert das Evaluierungsportfolio der ZPE.¹⁶ Auf flächendeckende Evaluierungen wird zugunsten einer Stichprobe verzichtet.

Durchgeführt wird eine Evaluierung durch ein externes Evaluierungsteam¹⁷, das von der GIZ bei Vorbereitung, Durchführung und Nachbereitung unterstützt wird. Aufgabe des Evaluierungsteams in der **Auftaktphase** ist es, alle wichtigen Aspekte des Projektes und die Perspektive der Beteiligten zu identifizieren, daraus die mit der Evaluierung zu beantwortenden Fragen zu entwickeln und die verfügbaren Daten zu prüfen. Die Auftaktphase umfasst eine Einführungsmission (i. d. R. vor Ort) und endet mit einem Auftaktbericht zum methodischen Vorgehen.¹⁸ Auf Grundlage der vorab festgelegten Methoden sammelt das Evaluierungsteam in der **Evaluierungsphase** weitere notwendige Daten. Neben Interviews mit den beteiligten Akteuren analysiert, bewertet und validiert es systematisch die Daten und verfasst einen Evaluierungsbericht.¹⁹

Das DEval führt seit dem Jahr 2020 eine Metaevaluierung²⁰ von Projektevaluierungen²¹ der deutschen EZ durch. Auch die GIZ will die ZPE zukünftig regelmäßig durch Metaevaluierungen überprüfen lassen.

¹⁶ Vgl. GIZ, Evaluierungsbericht 2017, S. 4.

¹⁷ Das Evaluierungsteam besteht in der Regel aus zwei Mitgliedern (internationale/r und lokale/r oder regionale/r Evaluator/Evaluatorin) und kann je nach Umfang und Komplexität der Evaluierung auch größer sein. Der internationale Evaluator/die internationale Evaluatorin wählt die weiteren Teammitglieder aus und benennt ihn/sie vor Auftragserteilung durch die GIZ.

¹⁸ Im Sprachgebrauch der GIZ wird der Auftaktbericht „inception report“ und der Evaluierungsbericht „final report“ genannt.

¹⁹ Näheres zum Vorgehen, vgl. GIZ, „*Das Evaluierungssystem der GIZ – Zentrale Projektevaluierung im BMZ-Geschäft*“.

²⁰ Metaevaluierung ist eine Evaluierung einer oder mehrerer anderer Evaluierungen zur Beschreibung und Bewertung ihrer Güte, Tauglichkeit und Bedeutsamkeit.

²¹ In diese Metaevaluierung sollen Projektevaluierungen der DO und zivilgesellschaftlicher Organisationen einbezogen werden, vgl. DEval-Evaluationen 2021 – 2023, Stand Januar 2021, Themenschwerpunkte, laufende und geplante Evaluierungen des DEval, http://www.deval.org/files/content/Dateien/Evaluierung/Policy%20Briefs/2021/DEval_Evaluierungsprogramm_2021-2023.pdf, letzter Zugriff am 26. Februar 2021.

3 Vorbereitung der Evaluierung

3.1 Stichprobenverfahren

(1) Im Verfahren der ZPE werden nicht alle Projekte evaluiert, sondern eine Stichprobe, die aus einer Grundgesamtheit gezogen wird. Die Grundgesamtheit umfasst alle in einem Jahr zu beendenden, von Ihnen mit mehr als 3 Mio. Euro finanzierten Projekte.²² Auch entsprechend kofinanzierte Projekte (zum Beispiel durch die Europäische Union -EU-) werden zukünftig Teil der Grundgesamtheit.²³ 95 % der zu evaluierenden Projekte werden durch die Stichprobe gewählt; die restlichen 5 % kommen frei nach bestimmten, vorher festgelegten Kriterien hinzu.²⁴

Ziel der stichprobenweisen Evaluierung ist, eine möglichst repräsentative Aussage darüber zu erhalten, ob die Projekte der Grundgesamtheit erfolgreich waren oder nicht (Erfolgsquote).²⁵ Für eine verlässliche Aussage muss die Stichprobe ausreichend groß sein. Die Stichprobengröße lag bei 41 % für die erste Grundgesamtheit, 48 % für die zweite Grundgesamtheit und 40 % für die dritte Grundgesamtheit²⁶.

²² Der Betrachtungszeitraum umfasst jeweils Projekte mit Ende-Soll Oktober des Folgejahres bis September des darauffolgenden Jahres.

²³ Hierzu hat sich die StbS nachträglich entschieden, da kofinanzierte Projekte mehr als die Hälfte der Projekte ausmachen. Mit der zuständigen Generaldirektion der EU will sich die StbS darüber abstimmen, wie Evaluierungen bestmöglich und zum beiderseitigen Nutzen koordiniert werden können. Kofinanzierte Projekte entfallen nur dann aus der Stichprobe, wenn der Geber der Kofinanzierung sie bereits vor nicht zu langer Zeit evaluiert hat und sie den qualitativen Anforderungen der GIZ entspricht.

²⁴ Diese Kriterien schlagen die operativen Bereiche, die StbS oder der Fach- und Methodenbereich der GIZ oder Ihr Haus vor.

²⁵ Die Erfolgsquote, z. B. 80 %, zeigt den Anteil der Projekte, die mit der Gesamtnote „erfolgreich“ bewertet wurden (zu den Einzelheiten der Bewertung der Projekte, siehe unter Tz. 4.4). Die Betrachtung der statistischen Repräsentativität bezieht sich daher auf die Abbildung dieser durchschnittlichen Erfolgsquote aller Projekte der Grundgesamtheit in der Stichprobe.

²⁶ Erste Grundgesamtheit: 108 von 265 Projekten mit Ende-Soll Oktober 2018-September 2019 wurden ausgewählt, vgl. Kurz-Papier der GIZ („one-pager“) vom 29. Juli 2019 zur Berechnung der Stichprobengröße für die Zentralen Projektevaluierungen (ZPE) der GIZ und Hinweise zur statistischen Repräsentativität.
Zweite Grundgesamtheit: 85 von 177 Projekten mit Ende-Soll Oktober 2019-September 2020 wurden ausgewählt, vgl. GIZ „one-pager“.

Inzwischen wurde eine Stichprobe für Projekte aus einer dritten Grundgesamtheit im Betrachtungszeitraum Ende-Soll Oktober 2020 – September 2021 aus 203 Projekten gezogen, die 40 % umfasst. Anders als in den vorhergehenden Stichproben wurde die Schichtung zunächst nach Finanzierungstiteln vorgenommen. Innerhalb des Titels Bilaterale TZ war eine zusätzliche regionale Schichtung möglich.

Die GIZ beabsichtigt, die Belastbarkeit der Stichproben Mitte 2021 zu überprüfen.²⁷

Die Ausgaben für rund 100 ZPE pro Jahr betragen knapp 8 Mio. Euro.

Zwei der von uns näher untersuchten ZPE waren EU-kofinanziert. Für diese beiden Projekte machte die GIZ die Vorgabe, in der Evaluierungsmethode auch den EU-Kontext zu berücksichtigen.

Neben der ZPE gibt es für alle Projekte der TZ weitere Überprüfungsmechanismen, u. a. durch jährliche Projektfortschrittsberichte mit Darstellung des aktuellen Stands der Zielerreichung.²⁸ Auch Projekte, die wegen eines Auftragswertes unter 3 Mio. Euro nicht in die Grundgesamtheit für die ZPE fallen, werden so überprüft.²⁹ Für bestimmte Anwendungsbereiche führen die Projektverantwortlichen vor Ort unter beratender Begleitung der StbS auch dezentrale Evaluierungen durch.³⁰

(2) Die GIZ hat bereits wesentliche Anstrengungen unternommen, um ein Stichprobenverfahren zu entwickeln, das eine repräsentative Aussage über den Erfolg aller Entwicklungsprojekte erlaubt. Wir erkennen auch an, dass die GIZ sich eine Anpassung oder Weiterentwicklung des Verfahrens nach der Überprüfung Mitte 2021 offenhält. Eine abschließende Bewertung, ob die Stichprobe mit dem derzeitigen Deckungsgrad und mit den zugrunde gelegten Parametern ein belastbares Ergebnis bringt, steht allerdings noch aus.

²⁷ Die Belastbarkeit der Stichprobe soll in Zukunft jeweils für zwei Jahre berechnet werden, entsprechend dem Vorgehen der KfW. Angewendet werden soll wie bisher schon ein mathematisch-statistisches Verfahren. Aufgrund von Laufzeitveränderungen oder der Vermeidung doppelter Evaluierungen (z. B. durch Kofinanzierungs-Geber) verändern sich sowohl Grundgesamtheit als auch Stichprobe laufend. Die tatsächliche Belastbarkeit der Stichprobe kann sicher erst nach Beendigung der letzten Evaluierung im Portfolio berechnet werden (hier: Mitte 2021). Abschätzungen sind zuvor möglich. Die StbS hat uns auf Anfrage schriftlich auch bestätigt: *„Die Erfolgsquote wird für jeweils zwei Stichproben gemeinsam, also alle 2 Jahre berechnet werden, sobald eine kritische Masse an Evaluierungen für diesen Zeitraum vorliegt. Anhand dieser Erfolgsquoten kann der Wert p ggf. angepasst werden, wenn dies nötig erscheint.“*

²⁸ Die GIZ prüft nach eigenen Angaben alle Projekte bei der Angebotserstellung. Bei der Durchführung der Projekte ist die GIZ auch dafür verantwortlich, laufend zu überprüfen, ob die erzielten Ergebnisse zum Projektziel führen („kontinuierliches wirkungsorientiertes Monitoring“). Schließlich erstattet sie regelmäßig Bericht. Laut der Handreichung 023 des BMZ berichten die DO dem BMZ jährlich. Die Berichte dienen der Rechenschaft und Information des BMZ zu wesentlichen Aspekten der Projektdurchführung und bilden ferner eine wichtige Grundlage für die politische Steuerung. Nach Beendigung der Maßnahme legt die GIZ einen Schlussbericht vor.

²⁹ Vgl. GIZ, *„Das Evaluierungssystem der GIZ – Zentrale Projektevaluierung im BMZ-Geschäft“*, S. 4.

³⁰ Vgl. GIZ, *„Das Evaluierungssystem der GIZ – Allgemeine Beschreibung“*, Schaubild, S. 8.

Angesichts der hohen Anzahl kofinanzierter Projekte halten wir es für geboten, sie in der Grundgesamtheit und in der Stichprobe zu belassen, damit diese repräsentativ bleibt. In einem zunehmend komplexeren Umfeld mit multilateralen Maßnahmen der EZ ist auch der Mehraufwand gerechtfertigt, sich mit den Gebern einer Kofinanzierung organisatorisch und inhaltlich zu Evaluierungen abzustimmen und den Evaluierungsauftrag ggf. anzupassen. Dies gilt zumindest dann, wenn hieraus spezielle Erkenntnisse über die Zusammenarbeit der Projektbeteiligten zum Nutzen von Projektkooperationen hervorgehen.

Da nicht mehr alle Projekte von der ZPE erfasst werden, könnten die Projekte unter 3 Mio. Euro BMZ-Finanzierung aus dem Blickfeld für eine abschließende Endbewertung geraten. Angesichts der Kosten einer Evaluierung von rund 80 000 Euro erscheint es zwar sachgerecht, diese Projekte nicht prioritär zu evaluieren. Allerdings sollte deren Wirksamkeit zusätzlich zu den allgemeinen Überprüfungsmechanismen überwacht werden.

(3) Wir haben empfohlen, dass die GIZ zu gegebener Zeit alle ZPE der kofinanzierten Projekte querschnittlich auswertet. So kann sie weitere Erkenntnisse über deren besondere Herausforderungen und Erfolgsfaktoren gewinnen. Die GIZ sollte sich, wie beabsichtigt, in geeigneten Abständen Gewissheit über die Belastbarkeit ihrer Stichproben für die ZPE verschaffen und die Belastbarkeit zwischen den Überprüfungen zumindest bestmöglich schätzen. Sie und die GIZ sollten Projektfortschritts- und -abschlussberichte kontrollieren und querschnittlich auswerten, besonders bei Projekten, die nicht der ZPE unterliegen.

(4) Sie haben erwidert, dass Sie unsere Empfehlung teilen, die kofinanzierten Projekte querschnittlich auszuwerten. Angesichts der hohen Anzahl und der Diversität von kofinanzierten Projekten hätten Sie und die GIZ ein besonderes Erkenntnisinteresse. Bezüglich der Belastbarkeit der Stichproben haben Sie vorgebracht, dass das Gesamtergebnis der ZPE durch die Festlegung der Größe der Stichprobe in keinem Fall vorbestimmt sei. Zentral für das ZPE-System sei vielmehr die Unabhängigkeit der Gutachter in jeder einzelnen ZPE. Die Größe der Stichprobe könne angepasst werden, sollte mehr oder weniger Variation in den Bewertungen zu finden sein. Eine Überprüfung der Belastbarkeit der Stichprobe werde alle zwei Jahre erfolgen. Außerdem haben Sie unsere Empfehlung geteilt, auch die Wirksamkeit der Projekte mit einer BMZ-Finanzierung unter 3 Mio. Euro ausreichend zu überprüfen. Solche Projekte sollen künftig von der StbS querschnittlich ausgewertet werden.

(5) Wir nehmen Ihre Ausführungen zur Kenntnis. Wir werden zu gegebener Zeit nachfragen, ob und welche Erkenntnisse Sie aus der Auswertung kofinanzierter Projekte gewonnen haben und zu welchen Ergebnissen die Abstimmung von Evaluierungen mit der EU geführt hat. Außerdem werden wir nachhalten, was die Belastbarkeitsprüfung der Stichprobe ergeben hat. Dazu gehört insbesondere, ob sich Änderungen im Bewertungssystem auf die Variation der Bewertungen und somit auf die Stichprobengröße ausgewirkt haben. Schließlich werden wir nachverfolgen, wie und in welchem Umfang die GIZ die Wirksamkeit von Projekten mit einer BMZ-Förderung unter 3 Mio. Euro kontrolliert.

3.2 Finanzierung der Evaluierungen

(1) Die Evaluierungen werden aus dem Budget des zu evaluierenden Projekts finanziert. Nach der jährlichen Festlegung der zu evaluierenden Projekte werden diese im GIZ-Intranet bekannt gegeben und die zuständigen AV informiert.³¹ Bei einem Viertel der von uns näher untersuchten ZPE betrug der Zeitraum zwischen der Information der AV und dem Projektende über ein Jahr; im Durchschnitt acht Monate.³² In einigen ZPE haben wir festgestellt, dass die AV die für die Evaluierung notwendigen Daten und Kostenblätter erst nach der Information zusammengestellt und ermittelt haben. In einem Projekt reichten die Projektmittel nicht mehr, um die Evaluierung zu finanzieren. Darauf verschoß die GIZ die Evaluierung um ein Jahr und finanzierte sie zentral. Die GIZ nannte uns als Grund für die frühzeitige Information, dass die AV die Finanzmittel für die Evaluierung in das verbleibende Projektbudget einplanen müssen.

(2) Wir haben Risiken in der frühzeitigen Information der AV gesehen: Für die AV besteht kein Anreiz, Projekte kontinuierlich zu steuern. Etwaige Lücken beim Projektmonitoring könnten Sie später noch „gefahrlos“ schließen, wenn ihr Projekt für die ZPE ausgewählt wurde. AV wiederum, deren Projekte nicht für die Evaluierung ausgewählt wurden, könnten weniger sorgfältig bei der restlichen Projektdurchführung sein. Diese Risiken ließen sich jedoch minimieren, wenn die ZPE über eine Umlage auf alle Projekte finanziert werden. Sicherlich benötigen die AV auch dann einen hinreichenden Vorlauf für die

³¹ GIZ, „FAQ Zentrale Projektevaluierungen“, Stand 20. April 2018.

³² Die Zeiträume zwischen Bekanntgabe der Stichprobe und Projektende betrugen 2,5; 4; 5; 6; 9; 9; 14 und 17 Monate. Ein Projekt lief noch; es fand eine Zwischenevaluierung statt.

organisatorischen und fachlichen Vorbereitungen der ZPE. Dieser müsste aber nicht mehrere Monate betragen, wenn die AV die Finanzmittel für die Evaluierung nicht mehr in das verbleibende Projektbudget einplanen müssten. Die vor Projektende stattfindende Auftaktphase der ZPE sollte auch für die AV zur Vorbereitung ausreichen und ihnen Gelegenheit bieten, ihre besonderen Anliegen und Belange in das Evaluierungskonzept einzubringen.

(3) Wir haben empfohlen, die ZPE über eine Umlage auf alle Projekte zu finanzieren. Nach der Umstellung sollten die AV erst kurzfristig (nicht früher als vier Wochen) über die anstehende Evaluierung informiert werden. Mängel beim regelmäßigen Monitoring und der Dokumentation der laufenden Projektdaten sollten im Evaluierungsbericht benannt werden, auch wenn sie nachträglich behoben wurden. Die GIZ sollte ihre Praxis überdenken, das Evaluierungsportfolio vor Durchführung der letzten ZPE der Stichprobe im Intranet zu veröffentlichen.

(4) Sie haben uns darüber informiert, dass Sie und die GIZ im September 2020 eine Vereinbarung über die Finanzierung der ZPE getroffen haben. Für die aus dem Haushaltstitel der bilateralen TZ finanzierten Projekte sei mit Geltung ab Oktober 2020 ein sog. Beitragsverfahren bei der GIZ eingerichtet worden.³³ Dabei trage jedes Projekt durch einen festen Beitrag von 40 % der voraussichtlichen Evaluierungskosten zur Gesamtfinanzierung bei. In den anderen Haushaltstiteln sei wegen zu geringer Anzahl evaluierbarer Projekte ein Beitragsverfahren nicht umsetzbar³⁴; dort bleibe es dabei, dass eine ZPE durch das jeweilige Projektbudget finanziert wird. Eine titelübergreifende Umlagefinanzierung widerspräche haushaltsrechtlichen Prinzipien, weil nicht mehr zugeordnet werden könnte, für welchen haushaltsmäßigen Zweck die Kosten der Evaluierung angefallen sind. Auf eine Veröffentlichung der geplanten Evaluierungen im Intranet vor Durchführung der Evaluierung könne in Zukunft verzichtet werden; die GIZ werde die AV der für die ZPE ausgewählten Projekte nach und nach informieren. Sie streben den von uns angeregten Zeitraum von

³³ Bei der Stichprobe 2019 der ZPE (Ende der Projekte Oktober 2020 – September 2021) waren 211 von 259 Projekten aus dem Titel bilaterale TZ (Kapitel 2301 Titel 896 03) finanziert. Weitere 11 Projekte wurden durch Kapitel 2301 Titel 896 06 - Internationale Zusammenarbeit mit Regionen für nachhaltige Entwicklung – und die übrigen 37 Projekte aus acht verschiedenen Titeln finanziert. Die Titel 896 03 und 896 06 aus dem Kapitel 23 01 sind gegenseitig deckungsfähig.

³⁴ Die geforderte Stichprobengröße von 38–42 % aller evaluierbaren Projekte könne bei einer geringeren Projektzahl als 20 nicht durchgängig eingehalten werden.

rund vier Wochen vor Beginn der jeweiligen ZPE an. Im Übrigen werden die angesprochenen Risiken bereits durch ein wirkungsorientiertes Monitoring für BMZ-finanzierte Projekte gemäß dem GIZ-internen Regelwerk minimiert.

(5) Wir begrüßen, dass Sie ein Beitragsverfahren für die Finanzierung der Evaluierungen eingerichtet haben und die Praxis für die Information über anstehende Evaluierungen gemäß unseren Empfehlungen anpassen werden. Aus unserer Sicht stehen einer titelübergreifenden Umlagefinanzierung keine haushaltsrechtlichen Bedenken entgegen, soweit darin Titel einbezogen werden, die untereinander deckungsfähig sind. Wir empfehlen daher, in das Beitragsverfahren auch die Projekte einzubeziehen, die durch den Haushaltstitel „Internationale Zusammenarbeit mit Regionen für nachhaltige Entwicklung“, der mit dem Haushaltstitel der bilateralen TZ deckungsfähig ist, finanziert werden.

4 Durchführung der Evaluierung

4.1 Ziel- und Ausgangswerte der Indikatoren

Nach Ihren „Leitlinien für die bilaterale Finanzielle und Technische Zusammenarbeit mit Kooperationspartnern der deutschen Entwicklungszusammenarbeit“³⁵ sollen die DO in ihrem Projektvorschlag für eine EZ-Maßnahme überprüfbare Ziele definieren. Für das Erreichen der Projektziele sollen bis zu fünf messbare Ergebnisse (*outputs*³⁶) definiert werden. Zur Messung der Ergebniserreichung werden dann drei bis fünf sog. SMART-Indikatoren³⁷ definiert und hierzu jeweils Ziel- und Ausgangswerte (*baselines*) angegeben. In den von uns untersuchten ZPE haben die Evaluierungen zu folgenden Erkenntnissen zu den Ziel- und Ausgangswerten der Indikatoren geführt:

- Die Indikatoren sind teilweise ungeeignet (unrealistische Zielwerte, nicht SMART, nur Output-bezogen).
- Die Ausgangswerte sind nicht vorhanden oder nicht ausreichend.
- Das kontinuierliche wirkungsorientierte Monitoring und die Datengewinnung bezogen sich nicht immer ausreichend und systematisch auf die festgesetzten Indikatoren.

³⁵ Vgl. Fn. 7.

³⁶ Outputs sind Leistungen, Produkte, Güter oder Kapazitäten.

³⁷ SMART-Indikatoren sind *specific* (spezifisch), *measurable* (messbar), *achievable* (erreichbar), *relevant* (relevant), *time-bound* (zeitlich festgelegt).

In Ansätzen hat die StbS darauf hingewirkt, die Planung und damit auch die Evaluierbarkeit künftiger Projekte zu verbessern³⁸, indem sie ihre Erkenntnisse für die Weiterentwicklung der Indikatoren mit dem GIZ- „Fach- und Methodenbereich“ abstimmte.³⁹ Sie beabsichtigt, die ZPE querschnittlich hinsichtlich der Planung und Durchführung von Projekten auszuwerten.⁴⁰

(2) Die Qualität des Projektvorschlags, insbesondere die darin definierten Ziel- und Ausgangswerte der Indikatoren, ist entscheidend dafür, wie gut sich Veränderungen während der Durchführung des Projekts im Vergleich zur Ausgangssituation beurteilen lassen.⁴¹ Die GIZ sollte geeignete Mittel finden, um Planung, Monitoring und Evaluierung besser zu synchronisieren. Bei Beteiligung mehrerer Partner⁴² würde sich bereits bei der Projektkonzeptionierung anbieten, ein gemeinsames Ziel-, Indikatoren- und Wirkungssystem zu entwickeln, um späteren Koordinierungsaufwand bei der Evaluierung zu vermindern.

(3) Wir haben empfohlen, dass die GIZ bereits bei der Erstellung der Projektvorschläge strenge Maßstäbe an die Bildung der Indikatoren anlegen sollte. Dabei haben Sie und die GIZ zu kontrollieren, ob Ihre Handreichungen für den Projektvorschlag so präzise befolgt werden, dass bei einer späteren Evaluierung tatsächlich valide Aussagen über den Projekterfolg möglich sind. Die StbS sollte ihre Erkenntnisse über die Evaluierbarkeit und Durchführung von Projekten an die AV und alle für die Prüfung, Planung und Steuerung von Projekten Beteiligten weitergeben. Über die Berichterstattung hinaus sollte das Monitoring als laufender systematischer Prozess gestaltet sein. Die Zielsetzungen und Aktivitäten anderer Partner sollten möglichst vollständig einbezogen (und nicht nur im Projektvorschlag beschrieben werden) und die Indikatoren und Wirkungsmodelle mit ihnen abgestimmt werden.

(4) Sie haben unsere Auffassung geteilt, dass Ziel- und Ausgangswerte von Indikatoren in Projektvorschlägen trotz verbindlicher Regeln und Vorgaben und der Qualitätsprüfung nicht immer hinreichend definiert worden sind. Die

³⁸ Zusammenwirken der Planungs-, Monitoring- und Evaluierungsphase.

³⁹ Im Papier der StbS, „Übersicht der Aggregationsindikatoren“, sind für jeden Indikator die Ansprechpersonen in der Abteilung Fach- und Methodenbereich genannt.

⁴⁰ Vgl. „Das Evaluierungssystem der GIZ – Allgemeine Beschreibung“, Ziffer 1.2.

⁴¹ Wir haben in verschiedenen Arbeitspapieren erkennen können, dass die StbS bereits an der Verbesserung der Indikatoren arbeitet.

⁴² Projektpartner oder Geber einer Kofinanzierung.

Auftragsbedingungen für die Prüfung neuer Projekte würden daher künftig in einem weiteren Schritt vorsehen, dass die Planung neuer Projekte auf Evidenzen⁴³ zu stützen ist. Sie haben in Abstimmung mit den DO eine Verfahrensinformation zu Qualitätsstandards von Indikatoren erarbeitet und im August 2020 in Kraft gesetzt, um die Qualität von Indikatoren zu verbessern. Sie wollen im Jahr 2021 im Rahmen der jährlichen Externen Qualitätskontrolle der GIZ prüfen, ob es bezüglich der Qualität der Indikatoren und beim wirkungsorientierten Monitoring weitere Anpassungsbedarfe gebe. Unsere Empfehlung zur verbesserten Synchronisierung von Planung, Monitoring und Evaluierung werden Sie zum Anlass für eine diesbezügliche Querschnittsauswertung aller ZPE durch die GIZ nehmen.

(5) Wir nehmen Ihre Ausführungen zur Kenntnis. Wir werden zu gegebener Zeit nachfragen, ob und welche Erkenntnisse zum Kooperationsmanagement Sie aus der Auswertung kofinanzierter Projekte gewonnen haben. Außerdem werden wir nachhalten, was die jährlichen externen Qualitätskontrollen der GIZ ergeben haben, insbesondere zum Handlungsbedarf bei der Qualität der Indikatoren und beim Monitoring.

4.2 Evaluierungsansätze

(1) Nach den Vorgaben der GIZ soll die bei der ZPE angewandte Evaluierungsmethode ermöglichen, die Wirkungen eines Projekts zu erfassen. Es soll eine plausible Kausalbeziehung zwischen Projekt und Wirkung hergestellt (Attribution) oder zumindest der Beitrag eines Projekts zu Wirkungen belegt werden (Assoziation). Es geht darum, „hinreichende Belege dafür zu sammeln, dass die beobachteten intendierten Wirkungen höchstwahrscheinlich auf das Projekt zurückzuführen sind.“⁴⁴ Die GIZ gibt den Evaluierungsteams einen theoriebasierten Evaluierungsansatz vor, um einen solchen sog. „robusten Wirkungs-

⁴³ Evidenz in der EZ liegt vor, wenn die entwicklungspolitischen Wirkungen im Partnerland durch belastbare Daten und Erfahrungen belegt werden können. Die hierfür erforderlichen Daten müssen mit wissenschaftlichen Methoden erhoben worden sein.

⁴⁴ Vgl. GIZ, „*Das Evaluierungssystem der GIZ – Zentrale Projektevaluierung im BMZ-Geschäft*“, S. 11.

nachweis“ führen zu können.⁴⁵ In einzelnen Fällen, an denen die GIZ ein besonderes Erkenntnisinteresse hatte, führte sie eine „Rigorous Impact Evaluation“ (RIE) durch. Damit soll ein gegenüber den sonstigen Methoden strenger Wirkungsnachweis geführt werden.⁴⁶ Die GIZ hat die Erfahrungen mit der RIE in einer Querschnittsanalyse ausgewertet. Eine RIE ist aufwendiger und kostenintensiver als ein theoriebasierter Ansatz. Sie stellt höhere Anforderungen an die StbS bei der Begleitung der Evaluierung. Außerdem müssen Projektplanung und Datenerhebung besonders gut entwickelt und abgestimmt sein, um eine spätere RIE vorzubereiten und zu ermöglichen. Bei allen von uns ausgewerteten ZPE haben die Evaluierungsteams einen theoriebasierten Evaluierungsansatz genutzt.⁴⁷

(2) Aus der Querschnittsauswertung zur RIE lässt sich ableiten, unter welchen Voraussetzungen ein robuster Wirkungsnachweis möglich ist: Je besser die Projektplanung und die laufende Datenerhebung während des Projektverlaufs, desto robuster kann in Zukunft der Wirkungszusammenhang nachgewiesen werden. Der zusätzliche Abstimmungsaufwand und die verstärkte Einbindung der StbS erspart nach unserer Ansicht in der Phase der Evaluierung Aufwand für die nachträgliche Rekonstruktion von Wirkungsmodellen und Indikatoren sowie für die nachträgliche Datenerhebung.

(3) Die GIZ sollte angesichts der Kosten jeder einzelnen ZPE darauf hinwirken, dass die Evaluierungen noch präzisere Aussagen über die Kausalbeziehung zwischen Projekt und Wirkung treffen können.

⁴⁵ Dazu gehören insbesondere die Verwendung des Wirkungsmodells aus dem Projektvorschlag, die Anwendung der „*Theory of Change*“ und die Kontributionsanalyse. Bei der „*Theory of Change*“ formulieren die Evaluierungsteams auf der Grundlage des Wirkungsmodells Wirkungshypothesen und ggf. Mechanismen der darin hinterlegten Ursache-Wirkungs-Beziehungen, die in der Evaluierung geprüft und bewertet werden können. Bei der Kontributionsanalyse erstellen die Evaluierungsteams eine Kontributionsgeschichte, die auf der Grundlage belastbarer, nachvollziehbarer und glaubhafter Evidenz die beobachteten Veränderungen und den Wirkungsbeitrag des untersuchten Projekts aufzeigt, vgl. GIZ, „*Das Evaluierungssystem der GIZ – Allgemeine Beschreibung*“, Ziffer 2.5.

⁴⁶ Z. B. (quasi-)experimenteller Ansatz mit Vergleichsgruppen, Identifizieren des sog. „Nettoeffekts“ durch Ausschluss anderer Ursachen, um festzustellen, ob und in wie weit eine Maßnahme eine Veränderung verursacht hat; vgl. auch: DEval, <https://www.deval.org/en/rigorous-impact-evaluation.html>.

⁴⁷ Dies war in der Regel eine Kombination von Kontributionsanalyse, *Theory of Change* und Nachvollziehen des Wirkungsmodells aus dem Projektvorschlag.

(4) Sie haben dargelegt, dass die GIZ mit einem theoriebasierten Ansatz (Kontributionsanalyse) arbeite, um Kausalbeziehungen plausibel zu belegen.⁴⁸ Die GIZ werde sich weiter engagieren, diesen methodischen Ansatz zu verbessern. Außerdem arbeite ein Thementeam daran, in der deutschen EZ zu einer systematischen und gegenstandsangemessenen Verankerung von RIE beizutragen. RIE seien allerdings aus methodischen Gründen nicht für alle wirksamkeitsrelevanten Fragestellungen geeignet und zielführend. Sie haben bekräftigt, dass das Verhältnis von Kosten und Aussagekraft der Evaluierung auch für Sie ein wichtiges Anliegen sei. Langfristig verfolge die StbS das Ziel, strategisch relevante Themen und Evidenzlücken unternehmensweit zu priorisieren. In drei von der EU kofinanzierten Projekten unterstütze die StbS aktuell die Koordination und Durchführung von RIE.

(5) Wir nehmen Ihre Ausführungen zur Kenntnis. Wir werden zu gegebener Zeit bezüglich der Fortentwicklung der Kontributionsanalyse und der Verankerung von RIE in der deutschen EZ nachfragen.

4.3 Effizienzbewertung

(1) Die GIZ betrachtet eine Entwicklungsmaßnahme als effizient, wenn mit den bereitgestellten finanziellen Mittel die maximalen Wirkungen der Entwicklungsmaßnahme erreicht werden.⁴⁹ Seit der Einführung der ZPE in 2017 bewertet die GIZ die Produktionseffizienz⁵⁰ einer Entwicklungsmaßnahme durch Zuordnung von Kosten zu Ergebnissen („Output“). Die GIZ wendet standardmäßig die Analysemethode „Follow-the-Money“ an, bei der im Nachgang eines

⁴⁸ Über die Kontributionsanalyse werde auf allen Wirkungsebenen (Output, Outcome und Impact) explizit der kausale Zusammenhang zwischen der evaluierten Maßnahme und den beobachteten Wirkungen analysiert und beschrieben (in einer Schätzung des Beitrags mit quantitativen Methoden oder in einer qualitativ dargelegten Wirkweise).

⁴⁹ Die Evaluierungsfragen der GIZ zur Effizienz lauten:

- Inwiefern hätten die Outputs bei gleichbleibender oder besserer Qualität mit dem gleichen Mitteleinsatz und unter den gleichen Rahmenbedingungen maximiert werden können?
- Inwiefern hätten Outputs durch andere Zuordnung der Mittel zwischen den Outputs maximiert werden können?
- Ob und wie wurden bei der Konzeption und im Verlauf der Umsetzung die Output-Mittel-Relation und Alternativen abgewogen?

vgl. GIZ, „Anleitung zur Effizienzbewertung“, S. 32 sowie S. 2 mit Verweis auf Palenberg, „Tools and Methods for Evaluating the Efficiency of Development Interventions“.

⁵⁰ Sie unterscheidet dabei die Produktions- und Allokationseffizienz. Produktionseffizienz bewertet die Transformation von Inputs zu Outputs, Allokationseffizienz von Inputs zur Outcome- und Impactebene. Die Allokationseffizienz wird gutachterlich durch die Evaluierungsteams eingeschätzt. Für die Betrachtung der Outputseite dienen die Wirkungsmatrix sowie der letzte Fortschrittsbericht mit der Zielerreichung als Datenbasis.

Projektes das Verhältnis der jeweiligen Kosten zu den entsprechenden Outputs bewertet sowie nach Potenzialen zur Ertragsmaximierung gesucht. Die GIZ hat ein Excel-Tool zur Erfassung der Kosten und deren Zuordnung zu den Outputs entwickelt.⁵¹

(2) Die Follow-the-Money-Methode und das Excel-Tool scheinen geeignet zu sein, die Kosten den Outputs eines Projekts nachträglich zuzuordnen. Es ist aus unserer Sicht zur Effizienzsteigerung und auch zur Steuerung der Projekte durch die AV darüber hinaus notwendig, die Kosten laufend zuzuordnen. Eine IT-Anwendung könnte den AV die für die Effizienzanalyse nötigen Daten (zu der Wirkungsmatrix, den Aktivitäten und den Kosten) kontinuierlich bereitstellen und wäre auch gegenüber Übertragungsfehlern weniger anfällig. Möglichst sollte sie auch Kofinanzierungen und Partnerbeiträge umfassen.

(3) Wir haben empfohlen, dass die GIZ die Kosten und Ergebnisse eines Projekts mittels einer einheitlichen und umfassenden IT-Anwendung laufend zuordnen sollte. Die IT-Anwendung sollte kostengünstig und mit wenig Aufwand für die AV behaftet sein, um deren Nutzungsbereitschaft zu erhöhen.

(4) Sie haben in Ihrer Stellungnahme ausgeführt, wie die Effizienz der Projekte sichergestellt und überprüft werden soll. Dabei unterscheiden Sie zwischen der Kostenkalkulation im Vorfeld des Projektes, der möglichst sparsamen Durchführung des Projekts und der Bewertung der Effizienz im Rahmen der Evaluierung. Bezüglich der (ex-post) Betrachtung der Effizienz werde dem Sparsamkeitsprinzip zukünftig besonders Rechnung getragen, indem die GIZ zusätzliche Evaluierungsfragen zur Effizienz aufnehme.⁵² Bei der Weiterentwicklung der IT strebe die GIZ an, den Stand der Kosten dem Stand der Ergebnisse (Output, Outcome) eines Projekts laufend IT-gestützt zuzuordnen. Durch diese Verknüpfung von Finanz- und Wirkungsmonitoring erhielten die AV damit künftig relevante Informationen zur wirtschaftlichen Steuerung der Projekte.

⁵¹ Dabei werden die Kosten der Zielerreichung einer Maßnahme auf Ebene der Indikatoren zugeordnet. Für die Zielerreichung der Indikatoren (Ist- und Zielwerte) müssen Zahlenwerte eingetragen werden. Auch die qualitativen Indikatoren eines Projekts sollen soweit wie möglich quantifiziert werden.

⁵² Diese zusätzliche Evaluierungsfrage der GIZ zur Effizienz lautet: „Inwieweit wurden die Inputs (finanziellen, personellen und materiellen Ressourcen) der Maßnahme im Verhältnis zu den erbrachten Outputs (Produkte, Investitionsgüter und Dienstleistungen) sparsam eingesetzt [...]? Ggf. als ergänzender Blickwinkel: Inwieweit hätten die Outputs (Produkte, Investitionsgüter und Dienstleistungen) der Maßnahme durch einen alternativen Einsatz von Inputs (finanziellen, personellen und materiellen Ressourcen) erhöht werden können [...]?", vgl zu den übrigen Fragen: Fn. 49.

(5) Wir erkennen die Anstrengung an, die Effizienz von Projekten bei der Planung, Durchführung und Abrechnung sicherzustellen. Wir begrüßen, dass bei der ex-post-Bewertung dem Sparsamkeitsprinzip beim Mitteleinsatz besonders Rechnung getragen werden soll. Der Evaluierung, aber auch der effizienten Projektdurchführung dient es sicherlich, dass die AV die Kostenzuordnung künftig IT-gestützt laufend vornehmen sollen. Wir regen an, zu gegebener Zeit querschnittlich festzustellen, ob sich die Bewertung des Kriteriums „Effizienz“ auf ein zuverlässiges und vollständiges Finanzmonitoring stützen kann und ob sich die Kosteneffizienz gesteigert hat.

4.4 Bewertung der Projekte

(1) Die Bewertungsskala umfasst insgesamt sechs Bewertungsstufen, drei im Bereich „erfolgreich“ und drei im Bereich „unbefriedigend“.⁵³ Die Gesamtbewertung setzt sich aus den jeweiligen Einzelnoten der bisherigen fünf DAC-Kriterien⁵⁴ zusammen, wobei diese zu je 20 % in die Gesamtbewertung einfließen.⁵⁵ In der Punkteskala entspricht eine Punktzahl von 67-100 einer der drei Bewertungsstufen „erfolgreich“ und von 0-66 Punkten einer der drei Bewertungsstufen „unbefriedigend“. Der Durchschnitt der Gesamtbewertungen der von uns näher untersuchten Projekte lag bei 81,5 Punkten („erfolgreich“).

⁵³ Der Notenbereich „erfolgreich“ ist aufgeteilt in die drei Stufen „sehr erfolgreich“, „erfolgreich“ und „eher erfolgreich“. Der Notenbereich „unbefriedigend“ ist aufgeteilt in die drei Stufen „eher unbefriedigend“, „unbefriedigend“, und „sehr unbefriedigend“.

⁵⁴ Das neu angenommene, sechste DAC-Kriterium „Kohärenz“ ist noch nicht berücksichtigt, siehe Fn. 13.

⁵⁵ Die von der StbS vorgegebene Evaluierungsmatrix gibt für jedes DAC-Kriterium maximal 100 Punkte vor. Nur für Unterkriterien (sog. „Bewertungsdimensionen“) innerhalb eines DAC-Kriteriums gibt es eine Gewichtung durch unterschiedliche Maximal-Punktzahlen.

Abbildung: Überblick über die Bewertungen der von uns näher untersuchten neun Projekte

100-Punkte-Skala	6 Bewertungsstufen	Gesamt-note	Relevanz	Nachhaltigkeit	Effektivität	Entwicklungs-politische Wirkungen	Effizienz
92-100	Stufe 1 = sehr erfolgreich				•	..	...
81-91	Stufe 2 = erfolgreich			•		..	..
67-80	Stufe 3 = eher erfolgreich	...	•		...		
50-66	Stufe 4 = eher unbefriedigend			..			
30-49	Stufe 5 = unbefriedigend						
0-29	Stufe 6 = sehr unbefriedigend						

Quelle: Bundesrechnungshof, eigene Darstellung; • = Häufigkeit der Bewertungen

Alle 24 evaluierten Projekte aus den Evaluierungs-Portfolios von Oktober 2017 bis September 2020⁵⁶ bekamen eine Gesamtnote im Notenbereich „erfolgreich“. Eines der Projekte erhielt die Gesamtnote „sehr erfolgreich“, 14 die Gesamtnote „erfolgreich“ und neun die Gesamtnote „eher erfolgreich“.⁵⁷

(2) Zwar halten wir das Vorgehen der GIZ für sachgerecht, nur Projekte im oberen Drittel der Punkteskala als „erfolgreich“ zu bezeichnen, um einen ausreichend hohen Qualitätsstandard zu gewährleisten. Allerdings zeigt unsere Auswertung der ZPE, dass die Evaluierungsteams immer zu einer Gesamtbewertung innerhalb des oberen Drittels der Punkteskala kamen. Auch die einzelnen DAC-Kriterien bewerteten sie überwiegend als „erfolgreich“. Die überdurchschnittlich guten Bewertungen der Relevanz glichen die eher unterdurchschnittlichen für die Nachhaltigkeit aus. Im Ergebnis erhielten dann alle Projekte eine Gesamtnote im Notenbereich „erfolgreich“. Eine genauere Aussage über den Projekterfolg lässt sich so nur aus der Punktzahl und aus dem Text der Evaluierung ablesen. Dies könnte auch darauf hinweisen, dass die Anforderungen an die Relevanz eines Projektes eher leichter und die der

⁵⁶ Nach Aussonderung von Projekten, die in zwei Portfolios geführt wurden, verblieben 24 Evaluierungen in den Evaluierungs-Portfolios von Oktober 2017 bis September 2020.

⁵⁷ Das entspricht 4 %, 58 % und 38 % der Projekte.

Nachhaltigkeit eher schwieriger zu erfüllen sind. Da sich das Kriterium der Relevanz eher auf das Projektkonzept bezieht, sollte es schon bei Annahme des Projektantrages befürwortet worden sein.

(3) Wir haben empfohlen, dass die GIZ ihr Bewertungssystem so verändern sollte, dass eine differenziertere und damit aussagekräftigere Gesamtbewertung erreicht wird. Dies schließe insbesondere die Bewertung der einzelnen DAC-Kriterien ein. Die GIZ könne erwägen, die DAC-Kriterien unterschiedlich zu gewichten. Dadurch könne die Betterbewertung für das Kriterium Relevanz ausgeglichen werden. Ferner könne dadurch die Bedeutung der Kriterien, die aus Ihrer entwicklungspolitischen Sicht besonders wichtig sind, auf die Gesamtbewertung eines Projektes erhöht werden.

(4) Sie haben erwidert, dass die Kriterien durch die Anwendung der neuen Fassung einer BMZ-Orientierungslinie⁵⁸ ab Oktober 2020 bei den staatlichen DO einheitlich behandelt würden. Effektivität, Impact und Nachhaltigkeit würden als sogenannte „K.O.-Kriterien“ verwendet: Wenn die Bewertung in einem dieser Kriterien „nicht erfolgreich“ sei, müsse das Projekt insgesamt als „nicht erfolgreich“ bewertet werden.⁵⁹ Das Relevanz-Kriterium werde damit niedriger gewichtet. Eine „sehr gute“ Bewertung der Relevanz könne die Gesamtnote dann nicht ausgleichen.

(5) Wir nehmen Ihre Ausführungen zur Kenntnis. Wir werden zu gegebener Zeit nachfragen, wie sich die Einführung von K.O.-Kriterien in den Bewertungen ausgewirkt hat.

4.5 Qualitätsprüfung durch die Stabsstelle

(1) Die StbS überprüft die Qualität der Auftakt- und Evaluierungsberichte. Bei diesen Qualitätschecks untersucht sie mittels einer Textanalyse und Auswertungstabelle, ob diese Berichte dem Evaluierungsstandard der methodischen Qualität (Genauigkeit) genügen. Mit einem Raster für die Qualitätsbeurteilung von Evaluierungsberichten will sie eine Mindestqualität der Evaluierung gewährleisten. Bei Nichterreichen dieser Mindestqualität bricht die GIZ die

⁵⁸ „Evaluierungskriterien für die deutsche bilaterale EZ“ zum Umgang mit den OECD-DAC-Evaluierungskriterien in Evaluierungen der deutschen bilateralen EZ.

⁵⁹ Die Einführung der K.O.-Kriterien Effektivität, Impact und Nachhaltigkeit sei per Beschluss in den Protokollen der AG-Evaluierung festgehalten und werde auch in die BMZ *Leitlinien Evaluierung* aufgenommen. Die Übertragung ins System der Zentralen Projektevaluierungen der GIZ sei erfolgt.

Evaluierung ab.⁶⁰ Die Qualitätschecks der GIZ wirken sich auch auf die Vergütung der Evaluierungsteams aus. Die Vergütung teilt sich auf in eine Vorauszahlung, Zwischenzahlung und Schlusszahlung. Liegt das Gesamtergebnis des Qualitätschecks beim Auftaktbericht unter 60 %, zahlt die GIZ lediglich die Vorauszahlung. Ist das Gesamtergebnis beim Evaluierungsbericht unter 60 %, beschränkt sich die Vergütung auf die Voraus- und Zwischenzahlung.

In Ihrem Haus haben Sie ebenfalls ein Prüfraster zur Qualitätsbeurteilung von Evaluierungsberichten aufgestellt. Hierüber fragen Sie nicht nur das Kriterium der Genauigkeit ab, sondern auch die beiden Kriterien der Nützlichkeit und Fairness.⁶¹ Zu den allgemein anerkannten Qualitäts-Evaluierungsstandards gehört neben den drei genannten Kriterien auch das Kriterium der Durchführbarkeit. Die Gesellschaft für Evaluation e.V. (DeGEval)⁶² hat diese vier Evaluierungsstandards entwickelt und definiert.⁶³ Anhand der Fairness und Durchführbarkeit wird die Prozessqualität gemessen.

(2) Nach unserer Auffassung ist es sinnvoll, bereits im laufenden Evaluierungsprozess auch Erkenntnisse über die Kriterien der Durchführbarkeit, Fairness und Nützlichkeit zu erhalten. Denn erst hierdurch kann zusammen mit dem Kriterium der Genauigkeit ein Gesamteindruck über die Qualität der Evaluierung entstehen. So verstehen wir auch das Prüfraster aus Ihrem Haus, das drei der anerkannten vier Evaluierungsstandards abfragt. Die Qualitätsüberprüfungen einschließlich der Methoden zur Datenerhebung sollten zugleich

⁶⁰ Der Zweck des Qualitätschecks besteht für die GIZ darin, eine Mindestqualität für die Fortsetzung des Evaluierungsprozesses zu gewährleisten. Eine Erfüllung sämtlicher Qualitätsstandards bedeutet eine Quote von 100 %. Bis zu 10 % weniger werden als „keine bis geringfügige Abweichung“ bewertet, bei 11 – 20 % als „leichte Abweichung“, bei 21 – 40 % als „deutliche Abweichung“ und über 40 % Abweichung als „schwerwiegende Abweichung“ der Vorgaben in Bezug auf ein Bewertungskriterium dar. Im letzten Fall bricht die GIZ die Evaluierung ab.

⁶¹ BMZ, Qualitätsstandards für Evaluierungsberichte und Raster für die Qualitätsbeurteilung von Evaluierungsberichten.

⁶² Die DeGEval ist ein Zusammenschluss von Forschungseinrichtungen, Hochschulen, Unternehmensberatungen, Politikberatungsunternehmen, Ministerien, Verwaltungen und Ressortforschungseinrichtungen zum Austausch über das Thema Evaluation.

⁶³ DeGEval, Standards für Evaluation, S. 10 ff.:

- Nützlichkeitsstandards: Evaluierung ist an den geklärten Evaluierungszwecken sowie am Informationsbedarf der vorgesehenen Nutzer und Nutzerinnen ausgerichtet.
- Durchführbarkeitsstandards: Evaluierung wird realistisch, gut durchdacht, diplomatisch und kostenbewusst geplant und ausgeführt.
- Fairnessstandards: In der Evaluierung wird respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen.
- Genauigkeitsstandards: Die Evaluierung vermittelt gültige Informationen und Ergebnisse zu Evaluierungsgegenstand und den Evaluierungsfragestellungen.

nicht zu kompliziert oder zeitaufwändig sein, sondern vielmehr zweckentsprechend maßvoll und praktikabel.

(3) Wir haben empfohlen, dass die GIZ die Qualitätschecks der Evaluierungen auf alle vier Evaluierungsstandards ausweitet und sich nicht allein auf die methodische Qualität beschränkt. Zur Vermeidung unverhältnismäßigen Durchführungsaufwandes könne es sich anbieten, dass die StbS ihre Einschätzungen zu den übrigen Evaluierungsstandards in den jeweiligen Evaluierungs-Abschlussberichten zusammenfasst.

(4) Sie haben dargelegt, dass die GIZ mit den AV und den Evaluierungsteams die Ansprüche zur Prozessqualität und Nützlichkeit in Vorgesprächen thematisiere und deren Einhaltung in der Auftakt- und Evaluierungsphase durch weitere Gespräche nachhalte. Konkrete Aspekte der Durchführbarkeit würden in der Auftaktphase verstärkt adressiert. Die Prozessqualität, aber auch die Nützlichkeit würden im Abschlussgespräch über Checklisten erfasst. Für die Bewertung der Nützlichkeit auf den verschiedenen Ebenen seien für ein systematisches Feedback Befragungen der AV und weiterer Beteiligter geplant. Zwischen- und Schlusszahlungen veranlasse die GIZ nur, wenn die folgenden Voraussetzungen vorliegen: Auftakt- bzw. Evaluierungsbericht erreichen mehr als 60 % methodische Qualität im Qualitätscheck und sind fristgerecht eingereicht worden. Zudem müssen die Prozessqualität sowie die in den Auftragsbedingungen hinterlegten DeGEval-Standards eingehalten worden sein.

(5) Wir begrüßen, dass die GIZ die Evaluierungen auch hinsichtlich ihrer Prozessqualität und Nützlichkeit verstärkt überprüfen will. Insoweit halten wir es für sinnvoll, schon vor und in der Auftaktphase die Abfragen zu formalisieren. Wie die Einhaltung der Zahlungsvoraussetzungen überprüft wird, sollte ebenfalls konkret und in formalisierter Form im Qualitätscheck verankert werden.

5 Nutzen von Evaluierungen

5.1 Nutzen auf Projektebene

(1) Die Evaluierungen sollen den Entscheidungsträgern Ihres Hauses, der GIZ und den Partnerländern ermöglichen, evidenzbasierte Entscheidungen zu treffen. Die StbS analysierte nach den ersten Erfahrungen mit der ZPE, wie sie den Nutzen von Evaluierungen für die verschiedenen Nutzungsebenen steigern

könne.⁶⁴ Sie machte Vorschläge, wie Evaluierungsberichte auf der Projektebene mit spezifischen, adressatengerechten und gut lesbaren Empfehlungen Entscheidungsfindungen unterstützen können.⁶⁵ Dazu gehörte auch, zu welchem Zeitpunkt eine Evaluierung Entscheidungen über das evaluierte Projekt, einem Folgeprojekt oder anderen Projekten in demselben Partnerland oder Sektor am meisten nützt. Eine Überlegung ist, bereits die Einführungsphase der Evaluierung für die Entscheidungsfindung über ein Folgeprojekt nutzbar zu machen und mit einer „Mini-Zwischenevaluierung“ zu ergänzen.⁶⁶

(2) Wir haben angemerkt, dass der Erfüllungsgrad des Evaluierungsstandards „Nützlichkeit“ auf der Projektebene sinken könnte, wenn die AV und sonstigen Beteiligten die Evaluierungsergebnisse nicht akzeptieren oder keinen Nutzen darin erkennen. Für die AV dürften besonders die operativen Erkenntnisse aus Evaluierungsprozess und Evaluierungsbericht wichtig sein. Daher dürfte der Nutzen des Evaluierungsberichts für die AV umso geringer sein, je später ein Projekt evaluiert wird. Für Entscheidungen über ein Folgeprojekt kann die GIZ nicht auf den Abschluss der Evaluierung warten, da ein Folgeprojekt dann in der Regel bereits läuft. Da sich aber die Evaluierungsteams in der Einführungsphase schon intensiv mit dem Projekt befassen, könnten ihre ersten objektiven Einschätzungen stärker für Entscheidungen über Förderungswürdigkeit und Erfolgsaussichten von Folgeprojekten im Partnerland genutzt werden.

(3) Wir haben die GIZ ermutigt, sich verstärkt dafür einsetzen, den Nutzen von Evaluierungen zu erhöhen. Ziel solle sein, dass die AV und übrige Beteiligte das Verhältnis von Nutzen zu Aufwand positiv bewerten. Die GIZ solle die Ergebnisse der Evaluierungen in die Entscheidung über das Zustandekommen eines Folgeprojekts und dessen Ausgestaltung einbeziehen.

(4) Sie haben mitgeteilt, dass Sie unsere Feststellungen teilen. Damit die Evaluierung für alle Projektbeteiligten nützlich ist, werde versucht, deren Erkenntnisinteressen bestmöglich zu berücksichtigen. Deren Einbindung während der Evaluierung solle auch Impulse für das laufende Projekt bieten. Falls es ein Folgeprojekt gebe, werde dieses in einem eigenen Kapitel in die Evaluierung einbezogen. Laut Rückmeldungen habe dies die Akzeptanz und Nützlichkeit

⁶⁴ Arbeitspapier der StbS, „Nützlichkeitsdimensionen der ZPE“, vom 13. Februar 2020.

⁶⁵ So sollen die Evaluierungsberichte den Entscheidungsträgern mit gesonderten Abschlussgesprächen am Ende der Evaluierung besser vermittelt werden.

⁶⁶ Arbeitspapier der StbS, „Nützlichkeitsdimensionen der ZPE“.

der ZPE für die AV erhöht. Auch die BMZ-Handreichung für Projektvorschläge sehe verbindlich vor, bei Folgeprojekten explizit auf Evaluierungsergebnisse und -empfehlungen Bezug zu nehmen. Die Umsetzung dieser Anforderung müsse aber in der Praxis noch weiter gefestigt werden. In der Auftaktphase einer Evaluierung solle nun verstärkt auf Schlussfolgerungen geachtet werden, die für Folgeprojekte relevant sind (wie die Qualität der Indikatoren).

(5) Wir begrüßen Ihre Anstrengungen, den Nutzen der Evaluierungen weiter zu erhöhen und diesen allen Projektbeteiligten zu vermitteln. Inwieweit sich dadurch die Qualität von Folgemaßnahmen oder neuen Projekten verbessern wird, lässt sich allerdings erst mittelfristig feststellen.

5.2 Nutzen oberhalb der Projektebene

(1) Die GIZ möchte aus den Evaluierungen auch über die Projektebene hinaus einen unternehmensinternen Nutzen ziehen. Hierfür hält sie einen Austausch ihrer verschiedenen Organisationseinheiten⁶⁷ und Hierarchieebenen für erforderlich. Mit Fach-, Diskussions- und Informationsveranstaltungen möchte sie die strategische Reflexion der Evaluierungen fördern. Außerdem stellt sie alle Evaluierungsberichte unternehmensintern elektronisch zur Verfügung. Die Initiativen der StbS beruhen auf der Erkenntnis, dass die Evaluierungen bisher nicht ausreichend genutzt wurden. In Anbetracht der Ausgaben für die ZPE stelle das Wissensmanagement und das Lernen aus Evaluierungen nach wie vor eine Herausforderung dar. Unternehmensintern würden Ergebnisse nur sehr aggregiert verwendet. Mögliche Erkenntnisgewinne aus den Projektevaluierungen würden so erst gar nicht ausgeschöpft.⁶⁸ Die StbS lässt Querschnittauswertungen⁶⁹ durchführen, um Hinweise für die Weiterentwicklung von Leistungsangeboten und von Sektor- oder Länderstrategien⁷⁰ zu erhalten.⁷¹ Ansätze für ein Empfehlungs-Monitoring (Umsetzungsfragebogen) sind vorhanden.⁷²

⁶⁷ Zu diesen Bereichen der GIZ gehören Fach- und Methodenbereich, Sektor- und Globalvorhaben, Länderbereiche und in den Partnerländern aktive operative Bereiche.

⁶⁸ Vgl. Protokoll über die Besprechung zum Austausch zu Lernen aus Evaluierungen.

⁶⁹ Laut Auskunft der StbS liegen neben Querschnittsauswertungen zu früheren dezentralen Projektevaluierung durch die Auftragsverantwortlichen solche zu folgenden Themen vor: Internationale Kompetenzentwicklung, Berufliche Bildung, Gesundheitsvorhaben, Bildungsvorhaben sowie Ländliche Entwicklung.

⁷⁰ Länderstrategien bilden den strategischen Rahmen für EZ. Sie beschreiben die Zusammenarbeit mit dem Partnerland, die Aktivitäten anderer Geber im Land sowie die Schwerpunkte des deutschen Engagements. Sie definieren Ziele und Handlungsfelder für diese Schwerpunkte; Quelle: Handreichung des BMZ Nummer HR 023.

⁷¹ Vgl. GIZ, „Das Evaluierungssystem der GIZ – Allgemeine Beschreibung“, Ziffer 1.2.

⁷² Undatierter Entwurf einer Evaluierungshandreichung, Nummer 3.2 f.

(2) Trotz der Aktivitäten der StbS, den fachlichen und operativen Unternehmensbereichen die Erkenntnisse aus den ZPE besser zu vermitteln, wird das Potential der ZPE nicht vollständig genutzt, um daraus strategische Rückschlüsse auf die TZ insgesamt zu ziehen. Angesichts der Kosten der ZPE besteht noch ein Missverhältnis von Kosten und Nutzen auf strategischer Ebene. Wir halten zudem ein Empfehlungs-Monitoring für notwendig. Es bietet sich an, dass die GIZ Wissensmanagement und Empfehlungs-Monitoring verknüpft.

(3) Wir haben empfohlen, dass sich über die StbS hinaus die gesamte GIZ-Leitungsebene dafür engagiert, dass gewonnenes Wissen genutzt und daraus gelernt wird. Dazu sollte die StbS zusätzlich zu ihrer nachfrageorientierten Vermittlung des Evaluierungswissens die Evaluierungsergebnisse noch gezielter in adressatengerechten Formen an unternehmensinterne Beteiligte weitergeben. Die GIZ sollte die wichtigsten Feststellungen und Empfehlungen der ZPE in eine Datenbank aufnehmen und thematisch auswerten. Dies ließe zumindest tendenzielle Erkenntnisse zu, die bei Bedarf in einer Querschnittsauswertung bestätigt werden könnten. Sie sollte insbesondere nachhalten, ob, inwieweit und in welcher Form die Adressaten den Empfehlungen nachkommen.

(4) Sie haben in Ihrer Stellungnahme ausgeführt, dass und wie die StbS die durch Evaluierungen gewonnenen Erkenntnisse über eine Vielzahl an Kanälen unternehmensintern weitergibt. Die StbS habe für die Jahre 2020 bis 2022 weitere Maßnahmen zur besseren Nutzung formuliert und mit deren Umsetzung begonnen. Dazu gehörten u. a. Evaluierungsergebnisse gemeinsam mit den fachlich zuständigen Einheiten sowie bestimmte Themen oder Sektoren querschnittlich auszuwerten. ZPE sollen außerdem in Zukunft als Bausteine von Programmevaluierungen genutzt werden und damit Evidenzen für die EZ insgesamt liefern. Die GIZ möchte die Evaluierungen mittels einer Software auswerten und die in den ZPE ausgesprochenen Empfehlungen nachhalten.⁷³ Darüber hinaus werde die GIZ ein geeignetes Verfahren der Nachverfolgung von Empfehlungen entwickeln und mit Ihnen abstimmen. Ihr Haus erhalte über einen mit der GIZ vereinbarten Verteiler⁷⁴ alle Evaluierungsergebnisse

⁷³ Eine Software zur Planung, Steuerung und für den Umgang mit Empfehlungen und Umsetzungsmaßnahmen werde der StbS ab 2022 im Rahmen der Einführung eines unternehmensweiten Audit-Management-IT-Tools zur Verfügung stehen.

⁷⁴ Im Verteiler seien das zuständige (Regional-)Referat, Referate, die für die jeweiligen Kennungen zuständig sind und das Evaluierungsreferat GS22.

(Hauptberichte und Kurzfassungen). Auf Wunsch organisiere die GIZ einen Austausch mit den Evaluierungsteams und Projektbeteiligten.⁷⁵

(5) Wir begrüßen, dass die GIZ die Verbreitung und Nutzung der Evaluierungen weiter verbessern und IT-gestützt die Empfehlungen aus den ZPE festhalten, auswerten und nachhalten will. Ihren Ausführungen konnten wir jedoch nicht entnehmen, ob und wie Sie im BMZ - von einzelnen Fällen abgesehen - die Evaluierungsberichte zu einzelnen Projekten behandeln. Wir empfehlen, die Erkenntnisse aus Querschnittsauswertungen systematisch bezüglich Schwerpunktthemen, Sektoren und Programmen für Ihre Grundsatzarbeit zu nutzen.

Reinert

Steinkamp

Beglaubigt: S. Erdmann, Tarifbeschäftigte

Wegen elektronischer Bearbeitung ohne Unterschrift und Dienstsiegelabdruck.

⁷⁵ Seit Einführung der ZPE hat die StbS auf Anfrage des BMZ den Austausch zu insgesamt sieben ZPE veranlasst. BMZ-Referenten haben bei einzelnen ZPE am Abschlussworkshop (Debriefing) teilgenommen.