



Grundsätze der Evaluierung

Ressortgemeinsame Evaluierungen

Orientierungslinien des BMZ

Inhalt

1	Vorbemerkung	3
2	Grundlagen	4
	Ziele	4
	Themen und deren Auswahl	4
	Standards und Kriterien	4
	Ressortübergreifende DEval-Evaluierungen	4
	Ressortgemeinsame Evaluierungen	5
3	Akteure – unterschiedliche Rollen für gemeinsamen Erkenntnisgewinn	6
	Steuerungsgruppe	6
	Referenzgruppe(n)	6
	Evaluierungsteam	7
	Partnerorganisationen	7
	Weitere Stakeholder	7
	Externes Review Panel	7
4	Evaluierungsprozess – Meilensteine zum Erfolg	8
4.1	Vorbereitung	8
4.1.1	Konzeption I: Terms of Reference (ToR)	8
4.1.2	Auswahl des Projektpartners und Bildung des Evaluierungsteams	9
4.2	Durchführung	9
4.2.1	Konzeption II	9
4.2.2	Datenerhebung	10
4.2.3	Fallstudien	10
4.3	Datenanalyse und Berichtslegung	11
4.3.1	Analyse von Fallstudien	11
4.3.2	Vorbereitung und Finalisierung des Abschlussberichts	11
4.3.3	Berichtsweg und Follow-up I	12
4.4	Qualitätssicherung über alle Phasen hinweg	12
4.5	Transparenz, Verbreitung und Follow-up II	12
4.5.1	Veröffentlichung und Verbreitung	12
4.5.2	Umsetzung von Empfehlungen	13
5	Literatur	14

1 Vorbemerkung

Um die Wirksamkeit des Regierungshandelns insgesamt im Sinne der Politikkohärenz für Entwicklung zu verbessern, strebt das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) in geeigneten Fällen Evaluierungen an, welche die ODA¹ mehrerer Ressorts umfassen. Dazu muss aufgrund des Ressortprinzips Einvernehmen zwischen den betroffenen Ressorts hergestellt werden.

Mit den Orientierungslinien für ressortgemeinsame Evaluierungen (OL Ressorts) soll dazu die Basis mit Blick auf grundlegende Prinzipien und Standards, Rollenverständnis und Verantwortlichkeiten sowie eine effektive Ablaufgestaltung gelegt werden. Sie beruhen auf international anerkannten Evaluierungsstandards und bewährter Praxis bei internationalen Gemeinschaftsevaluierungen sowie bei einer ressortgemeinsamen Evaluierung des Auswärtigen Amtes und des BMZ zur humanitären Hilfe im Ausland (2011; Kurzfassung 2012).

Auch wenn Variationen je nach Fallgestaltung (Untersuchungsgegenstand, Art der Beteiligten und Betroffenen, Hauptnutzenden etc.) erforderlich sein können, ist ein gemeinsames Verständnis zu den Eckpunkten, inklusive nachvollziehbarer Anpassungen, Voraussetzung für das Gelingen von ressortgemeinsamen Evaluierungen und hilft Transaktionskosten zu senken. Angesichts des vielfach generischen Charakters können die OL Ressorts auch als Orientierung für andere Gemeinschaftsevaluierungen dienen.

Im Lichte weiterer Praxis werden die OL Ressorts ggf. angepasst.

¹ Official development assistance (öffentliche Entwicklungsausgaben).

2 Grundlagen

Ziele: Bei ressortgemeinsamen Evaluierungen steht das Ziel im Vordergrund, die Wirksamkeit des Regierungshandelns insgesamt, über Ressortgrenzen hinweg, im Sinne der Politikkohärenz für Entwicklung (EU 2017) sowie des *Humanitarian-Development-Peace Nexus* (WHS 2016 und OECD DAC 2019) zu verbessern. Wie bei anderen Evaluierungen auch kann die Funktion in der Rechenschaftslegung oder im Erkenntnisgewinn für das Lernen zwecks zukünftiger Verbesserungen liegen. Der ressortübergreifende Ansatz bietet insbesondere ein hohes Potenzial, das gegenseitige Verständnis und das gemeinsame Lernen der beteiligten Ressorts durch die Erweiterung unterschiedlicher Perspektiven in gemeinsamen Handlungsfeldern zu befördern. Die breitere Grundlage bietet auch die Chance, die Legitimität einer Evaluierung und damit auch des Regierungshandelns zu erhöhen.

Themen und deren Auswahl: Evaluierungsthemen von ressortübergreifendem Interesse liegen in aller Regel auf der politisch-strategischen Ebene, sind also projektübergreifender Natur und beziehen sich meist auf Schnittstellen bzw. gemeinsame Handlungsfelder der Ressorts. Neben ggf. ressortspezifischen enthalten sie insbesondere ressortübergreifende Fragestellungen. Mit Blick auf das EU-weit beschlossene Kohärenzgebot kann auch Nicht-ODA-wirksames Regierungshandeln in oder mit Auswirkungen auf Entwicklungs- und Schwellenländer Gegenstand oder mitbehandelter Kontext von Evaluierungen mit BMZ-Beteiligung sein.

Die Auswahl von Evaluierungsvorhaben wird vorausschauend und ressortabgestimmt geplant. Voraussetzung für die Auswahl ist grundsätzlich die Evaluierbarkeit der Maßnahmen sowie die Wirtschaftlichkeit einer Evaluierung im Vergleich zu anderen Vorgehensweisen des Erkenntnisgewinns.

Standards und Kriterien: Ressortgemeinsame Evaluierungen orientieren sich als Ausgangspunkt an den politikfeldübergreifenden deutschen Standards für Evaluation (DeGEval 2016). Insbesondere bieten die etablierten OECD DAC-Prinzipien (1991), die Qualitätsstandards für Evaluierung (OECD DAC 2010) sowie internationale gute Praxis grundlegende Orientierung. Zentrale Prinzipien sind **Nützlichkeit** (und damit Partizipation der Betroffenen im Prozess, Transparenz, Verbreitung und Follow-up), **Unabhängigkeit** (insbesondere durch externe Durchführung sowie im Sinne von Unparteilichkeit und möglichst objektiven Bewertungen) und **Glaubwürdigkeit** (durch insbesondere Genauigkeit, hohe Qualität, Fairness und ethisches Verhalten). Im Falle von Maßnahmen in (auch potenziellen) Konfliktsituationen sind Evaluierungen konfliktensibel zu gestalten und ist insbesondere das „Do-no-harm-Prinzip“ zu beachten. Soweit konkrete Wirkungsuntersuchungen vorgesehen sind, sind diese nach Auswirkungen auf Frauen sowie benachteiligte Gruppen zu disaggregieren.

Als **Bewertungskriterien** kommen die etablierten OECD DAC-Evaluierungskriterien zur Anwendung, die ggf. für den jeweiligen Evaluierungsgegenstand oder -zweck (z. B. für humanitäre Hilfe, ALNAP 2006) angepasst oder auch ergänzt werden. Der Kohärenzaspekt – bei ressortübergreifenden Evaluierungen von besonderer Bedeutung – wird dabei als eigenständiges Bewertungskriterium² behandelt.

AUSGESTALTUNG

Ressortübergreifende DEval-Evaluierungen: Mit dem Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) verfügt der Bund über ein erfahrenes und zunehmend in Deutschland und international renommiertes Spezialinstitut, das wissenschaftlich fundierte Evaluierungen zu ODA-relevanten Maßnahmen unabhängig plant und durchführt – bisher für das BMZ-Portfolio. Die wirtschaftlichste und vom BMZ präferierte Form der Überprüfung ist daher, dass die betroffenen Ressorts auf Basis ihrer gemeinsamen Planung dem DEval vorschlagen – oder umgekehrt das DEval auf eigene Initiative den Ressorts vorschlägt – einen bestimmten Ausschnitt ihrer ODA-Portfolios ressortübergreifend zu evaluieren. Das Vorhaben geht im Falle der Zustimmung aller Beteiligten in das Evaluierungsprogramm des DEval ein. Die praxisbewährten, auf Basis der zuvor genannten Prinzipien und Standards entwickelten DEval-Standards und Verfahrensabläufe

² Seit Dezember 2019 gilt Kohärenz als eigenständiges DAC-Evaluierungskriterium, vgl. OECD DAC 2019.

sichern Best Practice bei Konzeption, Durchführung und Prozessgestaltung. Sie werden im Detail angepasst, insbesondere um den beteiligten Ressorts gleiche Informations- und Kommentierungsrechte wie dem BMZ einzuräumen. Eine zwischen den Ressorts zu vereinbarende Nutzung der DEval-Kapazitäten hilft Transaktionskosten der Ressorts zu senken (keine Ausschreibung erforderlich; geringerer Qualitätssicherungs- und Prozessgestaltungsaufwand) bei nachgewiesener methodischer Qualität. Eventuell fehlendes fachliches Know-how zum Zuständigkeitsbereich anderer Ressorts kann durch Untervertragnahme spezialisierter Anbieter bzw. von Fachpersonal ausgeglichen werden.

Die anderen Ressorts beteiligen sich im Falle eines erheblichen Anteils am Evaluierungsgegenstand möglichst finanziell durch eine Zuwendung.

Ressortgemeinsame Evaluierungen: Soweit zwischen den Ressorts keine Einigung über eine ressortübergreifende Evaluierung durch das DEval erzielt werden kann, können dem Ressortprinzip entsprechend, um Zielsetzungen und Bedürfnisse aller beteiligten Ressorts widerzuspiegeln, auch verschiedene Optionen³ *ressortgemeinsamer* Evaluierungen Anwendung finden. Soweit mehrere Ressorts in erheblichem Maße betroffen sind, empfiehlt sich, zumal ohne ausreichende Erfahrungswerte, ein hohes Maß an „gemeinsam“. Dafür spricht auch, dass bereits die Themenauswahl bei ressortgemeinsamen Evaluierungen meist impliziert, dass die Ressorts gleichermaßen oder jeweils in relevantem Ausmaß betroffen sind.

Der wesentliche Unterschied zu *ressortübergreifenden* DEval-Evaluierungen liegt darin, dass die Ressorts hierbei die Prozesskoordination und andere, sich aus der Natur der Sache ergebende Aufgaben übernehmen, unter Wahrung der zuvor genannten Prinzipien und Standards. Soweit zeitlich verfügbar und dazu bereit, übernimmt das DEval die Durchführung im Rahmen einer kooperativen⁴ Lösung für das „Projekt“ Evaluierung. Der Projektpartner (Firma, Institut) wird durch ein wettbewerbliches Verfahren des auftragserteilenden Ressorts ausgesucht. Lediglich unter außergewöhnlichen Umständen kommt auch eine Durchführung nur durch kommerzielle Anbieter oder andere Institute in Betracht.

Die folgende Beschreibung von Rollen und Aufgaben der Akteure sowie zum Evaluierungsprozess (Kapitel 3 und 4) behandelt diese zweite Option, ressortgemeinsame Evaluierungen.

³ Grundsätzlich ist eine große Bandbreite von „alles gemeinsam“ bis hin zu einer „delegierten Kooperation“ (wobei ein Ressort weitgehend nur als „stiller Partner“ fungiert) möglich. Ein hohes Maß an Delegation, verbunden z. B. mit einer nur selektiven Kommentierung von Evaluierungsprodukten oder einer Teilnahme an der Referenz- und nicht der Steuerungsgruppe, wäre beispielsweise sinnvoll, wenn das Portfolio des fraglichen Ressorts nur in geringem Umfang betroffen ist.

⁴ Der Begriff „kooperativ“ wird hier verwendet, um unterschiedliche Formen der Zusammenarbeit, ohne Festlegung der Rechtsform, abzudecken.

3 Akteure – unterschiedliche Rollen für gemeinsamen Erkenntnisgewinn

Gutes Evaluierungsmanagement und Rollenklarheit unterstützen die Qualität, die Unabhängigkeit (und damit Glaubwürdigkeit) sowie die Nützlichkeit von Evaluierungen. Auch wenn Variationen möglich sind, hat sich die im Folgenden beschriebene Funktionsaufteilung bewährt.

Alle Beteiligten wahren die Vertraulichkeit während des gesamten Ablaufs einer Evaluierung.

Steuerungsgruppe (SG): Die jeweiligen **Evaluierungseinheiten oder -beauftragten (Evaluierungsstellen)** der hauptverantwortlichen Ressorts bilden eine interministerielle SG. Die Hauptfunktion liegt in der Prozesssteuerung, der Sicherung der Unabhängigkeit sowie der Unterstützung der Qualität der Evaluierung, ggf. mit externer Unterstützung⁵.

Die SG trifft zu allen Meilensteinen abgestimmte Entscheidungen, unbeschadet einer auch arbeitsteiligen Vorgehensweise. Sie ist verantwortlich für die Koordination zwischen den Ressorts und für die Kommunikation innerhalb der vertretenen Ressorts. Sie sorgt für die Beteiligung relevanter Akteure entlang des gesamten Prozesses und übernimmt insofern die Prozesssteuerung, die den gesamten Zyklus von der Planung bis zur Verbreitung und zum Follow-up umfasst (vgl. 4). Zu den operativen Aufgaben gehört insbesondere die eigenständige Erstellung der Leistungsbeschreibung (Terms of Reference, ToR) und die Auswahl einer geeigneten Firma auf Basis der anvisierten kooperativen Lösung (unter Federführung des beauftragenden Ressorts). Damit ist sie auch Hauptansprechpartner für das Evaluierungsteam. Zu weiteren Aufgaben vgl. 4. Jeweils zu Beginn einer Evaluierung richtet die SG eine Referenzgruppe ein.

Referenzgruppe (RG): Die interministerielle RG besteht zumindest aus den thematisch federführenden Referaten und weiteren fachlich einschlägigen Arbeitseinheiten der Ressorts. Die Beteiligung von Mittler-, Träger- und Partnerorganisationen, ggf. weiterer Stakeholder und Externer, empfiehlt sich je nach Fragestellung (vgl. unten). Bei länderbezogenen Evaluierungen sollten in jedem Fall Vertretungen der deutschen Botschaft, insbesondere WZ-Referentinnen oder -Referenten, sowie – je nach Fragestellung – weitere Fachvertretungen in Kommentierungen einbezogen werden. Die Hauptfunktion der RG ist es, „Sounding Board“ die thematisch-fachliche Genauigkeit einer Evaluierung zu fördern. Gleichzeitig ermöglicht die durchgängige Beteiligung an Meilensteinen der Evaluierung den Mitgliedern das Lernen im Prozess und unterstützt damit die Nützlichkeit.

Ggf. bilden einzelne Ressorts zusätzlich eine ressortspezifische RG, soweit dies zu spezifischen Zielen der Evaluierung, beispielsweise mit Blick auf Besonderheiten der Implementierungsstruktur, beiträgt.

Eine besondere Rolle in der RG spielen die für den jeweiligen Evaluierungsgegenstand **thematisch federführenden Referate** der Ressorts. Sie sind die Hauptadressaten und intendierten Nutzenden von Evaluierungen. Ihr Erkenntnisinteresse steht damit im Vordergrund. Nach Abschluss einer Evaluierung sind sie für die Umsetzung (akzeptierter) Empfehlungen verantwortlich.

Die RG wird von der SG zu allen wesentlichen Meilensteinen von Evaluierungen eingeladen und periodisch über den Stand der Evaluierung informiert. Die Mitglieder der RG informieren und beraten die Evaluierungseinheiten und das Evaluierungsteam fachlich während des gesamten Prozesses, geben Auskünfte, vermitteln Kontakte, stellen erforderliche Daten und Dokumente bereit, informieren aktiv über aktuelle Entwicklungen

⁵ Soweit ein anderes Ressort über keine eigenständige Evaluierungsstelle verfügt, müssen Sonderregelungen getroffen werden, wie beispielsweise bei der Evaluierung „Humanitäre Hilfe im Ausland“ praktiziert.

des Evaluierungsgegenstandes und ggf. des Kontextes; sie kommentieren Zwischenprodukte und Berichtsentwürfe mit Blick auf die sachliche Korrektheit. Die letztendliche Verantwortung für die Ergebnisse der Evaluation liegt jedoch bei der Leitung des Evaluierungsteams.

Evaluierungsteam: Das Evaluierungsteam inklusive dessen Leitung wird von den Projektpartnern unter Berücksichtigung unterschiedlicher Qualifikationen mit Blick auf die Aufgabenstellung ausgewählt. Es besteht aus Männern und Frauen und umfasst möglichst auch lokale oder regionale, qualifizierte Evaluatorinnen und Evaluatoren. Die Hauptfunktion liegt in der Durchführung der Evaluation gemäß den anerkannten Qualitätsstandards. Mögliche Interessenkonflikte werden vor der Auswahl offengelegt. Das Evaluierungsteam erhält Zugang zu allen einschlägigen Daten und Dokumenten der betroffenen Bundesministerien sowie, soweit erforderlich, der Mittlerorganisationen (Durchführungs- oder Förderorganisationen) unter Wahrung der Vertraulichkeit und des Datenschutzes, verbunden mit Löschpflichten.

Partnerorganisationen: Die zu evaluierenden Maßnahmen werden in der Regel von oder zumindest mit lokalen Partnern implementiert. Bei der Planung, Durchführung und Berichtslegung von Evaluationen werden daher auch Verantwortliche aus Partnerlandorganisationen so weit wie möglich und dem Zweck angemessen einbezogen. Die Bandbreite reicht von einer Beteiligung in Form einer Gemeinschaftsevaluation (Vertretung in der SG) über die Einrichtung einer „Referenzgruppe Süd“, die zu allen Meilensteinen beteiligt wird, bis hin zur bloßen Information von Partnerorganisationen. Soweit der Evaluierungsgegenstand auf Regierungsabkommen oder Verträgen mit Regierungsstellen beruht, sollten Regierungsstellen zumal vor Feldmissionen über die Botschaft zumindest vorab informiert werden.

Weitere Stakeholder: Auch Mittler- oder Trägerorganisationen spielen vielfach eine bedeutsame Rolle bei der Durchführung der zu evaluierenden Maßnahmen. Diese können internationale Organisationen, Durchführungs- oder zivilgesellschaftliche Organisationen sein. Ihr fachliches Wissen und ihre Kenntnisse des Kontextes sollten durch eine Beteiligung an der Referenzgruppe genutzt werden, insbesondere dann, wenn sich abzeichnet, dass sie auch Adressaten von Empfehlungen sein werden. Da in aller Regel Partnerorganisationen und lokale Vertretungen von Mittler- und Trägerorganisationen im Zuge von Feldmissionen interviewt werden, werden diese zumindest im Rahmen eines Workshops zum Abschluss einer Feldmission beteiligt.

Einen Sonderfall stellen **Gemeinschaftsfinanzierungen**, an denen weitere Geber beteiligt sind, sowie **zweckgebundene Beiträge an internationale Organisationen** dar. In diesem Fall muss den anderen Gebern bzw. der internationalen Organisation eine angemessene Beteiligung angeboten und auf Wunsch eingeräumt werden.

Externes Review Panel: Je nachdem ob und wenn welches Modell der externen Qualitätssicherung gewählt wird, können ein oder mehrere Externe mit thematischer und/oder evaluatorischer Qualifikation als eigenständiges Gremium oder in die RG integriert vorgesehen werden. Zu möglichen Funktionen vgl. im Einzelnen 4.4.

4 Evaluierungsprozess – Meilensteine zum Erfolg

Vor Beginn des konkreten Evaluierungsprozesses muss Einigkeit zwischen den Ressorts – über ihre Evaluierungsstellen – zu den Eckpunkten hergestellt werden. Dazu gehören a) Ziel(e) der Evaluierung, (Vor-)Auswahl des Evaluierungsgegenstandes unter Berücksichtigung der Evaluierbarkeit sowie die Anwendung grundlegender Prinzipien und Standards und die Ausgestaltung (vgl. 2. Grundlagen); b) eine grundsätzliche Rollen- und Aufgabenklärung (vgl. 3. Akteure) sowie c) eine Einigung über die wesentlichen Ablaufschritte und Produkte (vgl. im Folgenden). Auf dieser Basis wird die grundsätzliche Verfügbarkeit des DEval für eine kooperative Lösung geklärt. Vor Beginn wird ebenfalls geklärt, ob weitere Ressorts betroffen sind und eingeladen werden sollen, an der Steuerungsgruppe oder der Referenzgruppe teilzunehmen.

Der im Folgenden skizzierte Ablauf hat sich bei Evaluierungen vielfach bewährt, unbeschadet sachgerechter Variationen im Einzelnen. Er ist sequenziell nach Phasen gegliedert und nennt zentrale Produkte und Meilensteine für Referenzgruppentreffen. Unverzichtbare Produkte sind die Leistungsbeschreibung (Terms of Reference, ToR), der „Inception Report“, eine Zusammenfassung der Ergebnisse, erste Schlussfolgerungen und Empfehlungen (meist ausführliche Präsentation) sowie der Abschlussbericht (erst Entwurf, dann final).

Referenzgruppentreffen werden von der SG auf längere Sicht geplant. Die zu besprechenden Unterlagen werden mindestens zwei Wochen, der Entwurf des Abschlussberichts mindestens drei Wochen vor den Treffen zur Verfügung gestellt.

4.1 VORBEREITUNG

Produkte: Leistungsbeschreibung (ToR), Verträge (i. d. R. Dienstleistungsvertrag bzw. Zuwendungsvertrag und Kooperationsvertrag) sowie formlose Vereinbarung

Austausch: ein oder mehrere Treffen mit den relevanten Fachreferaten der Ressorts und schließlich der (ggf. erweiterten) Referenzgruppe zu den ToR

4.1.1 Konzeption I: Terms of Reference (ToR): Die SG erstellt im Austausch mit den thematisch federführenden Referaten, anderen fachlich einschlägigen Referaten der Ressorts und ggf. den Mitgliedern einer umfassenderen Referenzgruppe eigenverantwortlich die Terms of Reference (ToR). In einem *ersten Schritt* werden insbesondere Ziel(e), Zweck und intendierte Nutzung geklärt. Das Erkenntnisinteresse der intendierten Nutzenden mit Blick auf die Portfoliosteuerung ist leitend für die Evaluierungsfragen, soweit diese mit den verfügbaren Ressourcen in Einklang gebracht werden können. Sie werden von der SG abgeglichen mit Blick auf die ressortübergreifende Steuerungszielsetzung und ggf. weitere Zielsetzungen der Evaluierung (z. B. Rechenschaftslegung aufgrund von Nachfragen aus dem Deutschen Bundestag oder Monita des BRH) und ggf. ergänzt.

Darauf aufbauend entwickelt die SG die → **Leistungsbeschreibung (ToR)**. Diese enthält Hintergrund, Anlass, Ziele, Gegenstand inklusive Zeitraum, Aufgaben, ggf. methodische Hinweise, Produkte (s. auch Verbreitung, Umsetzungsplanung) sowie einen vorläufigen Zeitplan und das Anforderungsprofil des Gutachterteams; sie skizziert die Governance-Struktur und die angestrebte Durchführung zusammen mit DEval, verweist auf die erwarteten Qualitätsstandards und gibt einführende Hinweise auf Literatur und Dokumente.⁶ Dabei sollte insbesondere (zusammengefasst) das *ressortübergreifende* Erkenntnisinteresse, vorzugsweise zusätzlich in Form von *kriterienspezifischen* Evaluierungsfragen, deutlich werden. *Ressortspezifische* Fragen können diese

⁶ Die Notwendigkeit weiterer Angaben kann sich aus den Vergabeanforderungen ergeben.

ergänzen. Zeitplan und finanzielle Ressourcen werden so bemessen, dass die Ziele der Evaluierung erreicht werden können: Aufgabenstellung und die Ressourcen müssen im Einklang stehen und die Zeitplanung berücksichtigt die intendierte Nutzung.

Die ToR lassen Raum für konzeptionelle Weiterentwicklung im Rahmen des Inception Reports, in Abwägung des übergreifenden Erkenntnisinteresses und der Umsetzbarkeit im Einzelnen (z. B. mit Blick auf Datenverfügbarkeit, Zugänglichkeit/Sicherheit, Wirtschaftlichkeit des Ressourceneinsatzes inkl. Zeitdauer).

Da das DEval auf Basis der ToR entscheidet, ob es im Rahmen einer kooperativen Lösung zur Verfügung steht, wird es rechtzeitig bei der Erstellung der ToR konsultiert, um einen möglichst friktionslosen weiteren Ablauf zu gewährleisten.

4.1.2 Auswahl des Projektpartners und Bildung des Evaluierungsteams: Der Partner für das Evaluierungsprojekt (Firma, Institut)⁷ wird mittels der ToR durch einen transparenten und wettbewerblichen Auswahlprozess – in der Regel durch eine Ausschreibung^{8, 9} – mit Verweis auf die gewünschte Durchführung in Zusammenarbeit mit DEval identifiziert und beauftragt. Um die Unabhängigkeit zu gewährleisten, wird sichergestellt, dass die Teammitglieder zuvor weder an der Planung noch an der Durchführung der zu evaluierenden Entwicklungsmaßnahme in nennenswertem Umfang beteiligt waren und dass keine sonstigen Interessenkonflikte vorliegen. Bei der Auswahl des Angebots werden Evaluierungs- und Fachkenntnisse (Kenntnisse und Erfahrungen mit vergleichbaren Evaluierungsgegenständen) berücksichtigt. Bei umfangreichen, strategischen Evaluierungen haben sich bewährt: a) ein zweistufiges Verfahren (1. Interessenbekundung und 2. Aufforderung zur Abgabe eines Angebots) und b) ein Interview mit den präqualifizierten Bietern zum besseren Verständnis ihrer Angebote.

Nach Beendigung des arbeitsteilig durchgeführten und gleichzeitig abgestimmten Wettbewerbs schließt das hierzu federführende Ressort einen → **Dienstleistungsvertrag (oder Zuwendungsvertrag)**.

Die Beteiligung des DEval erfolgt in Form einer **formlosen Vereinbarung** zwischen dem BMZ und dem DEval auf Basis der ressortgemeinsamen ToR.

Das DEval und der Projektpartner schließen auf dieser Basis, unter Beteiligung der Ressorts, einen → **Kooperationsvertrag**, mit dem die Form der Zusammenarbeit mit Rechten und Pflichten (z. B. Informations- und Datenaustausch), die finale Aufgabenverteilung und die Teamzusammensetzung inklusive -leitung aufgabengerecht geregelt wird.¹⁰

4.2 DURCHFÜHRUNG

Produkte: „Inception Report“; „Debriefing Note(s)“

Austausch: SG- und ggf. RG-Treffen zum Auftakt; RG-Treffen zum „Inception Report“

4.2.1 Konzeption II: Bewährt hat sich ein **Auftakttreffen** der SG, ggf. gemeinsam mit der RG, mit dem Evaluierungsteam, um ein gemeinsames Verständnis von Zielen, Untersuchungsgegenstand und zentralen Evaluierungsfragen sicherzustellen. Dabei kann das Evaluierungsteam eigene Vorschläge einbringen und verweist ggf. auf Grenzen der Durchführbarkeit. Als gute Praxis hat sich erwiesen, dass das Evaluierungsteam mit der SG die ToR sachgerecht bis zum „Inception Report“ weiterentwickelt.

⁷ Sollte keine Einigung über eine Beteiligung des DEval zustande kommen, muss die gesamte Evaluierung ausgeschrieben werden.

⁸ Als Alternative zu einem Auftrag käme, z. B. im Falle einer eher wissenschaftlichen Begleitforschung mit evaluativem Charakter, auch eine Zuwendung infrage. Zu beachten: Die Finanzierungsform hat Folgen für die Eigentumsrechte an den Produkten.

⁹ Eine weitere Option wäre, dass das DEval Gutachterinnen oder Gutachter als Mitglieder des Evaluierungsteams unter Vertrag nimmt oder mit einem anderen Institut zusammenarbeitet, das über spezielles Know-how zu Maßnahmen und Politiken der oder des anderen Ressorts verfügt.

¹⁰ Einzelheiten der Vorgehensweise hängen insbesondere auch von Vergabebestimmungen ab und sind daher nicht Gegenstand der Behandlung dieser OL Ressorts.

Auf Basis der von den Ressorts und ggf. von Mittlerorganisationen bereitgestellten Daten und Dokumente sowie ersten Interviews erstellt das Evaluierungsteam einen → „**Inception Report**“ (IR): Basierend auf einer vollständigen Übersicht aller einschlägigen Maßnahmen enthält der IR eine ggf. vorläufige qualitative und quantitative Portfolio- und eine Kontextanalyse, die spätestens im Abschlussbericht durch eine Literaturlauswertung ergänzt wird. Wesentlicher Bestandteil des IR ist der Abgleich der ggf. unterschiedlichen Wirkungslogiken der Maßnahmen der Bundesministerien und die Identifizierung der Schnittstellen und Anknüpfungspunkte im Hinblick auf die in den ToR formulierten gemeinsamen bzw. ressortübergreifenden Erkenntnisinteressen¹¹. Im „Inception Report“ werden die finalen Untersuchungsfragen, das Analyseraster (Evaluierungsmatrix), der methodische Ansatz (Design und Methoden) sowie das weitere Vorgehen inkl. Zeitplan und Produkte im Einzelnen (z. B. auch zu Fallstudien, deren Auswahl auf nachvollziehbaren Kriterien beruht; vgl. dazu auch 4.2.3 und 4.3.1) festgelegt.

Der Entwurf des „Inception Reports“ wird in einer **RG-Sitzung** besprochen. Die SG fasst die wesentlichen Punkte zusammen (durch Protokoll oder Textkommentare zumal zu sachlichen Richtigstellungen). Das Evaluierungsteam erläutert textlich oder mittels eines Antwortrasters, welche Hinweise es insbesondere nicht – und warum nicht – übernommen hat.

Bei sehr komplexen Vorhaben, die mehrere Fallstudien mit Erhebungen zumal in unterschiedlichen Ländern erfordern, kann auch eine explorative Fallstudie vor Erstellung des „Inception Report“ oder eine Pilotfallstudie danach erforderlich sein.

4.2.2 Datenerhebung: Das Evaluierungsteam setzt seine Datenerhebung ggf. in Deutschland und bei anderen Gebern sowie insbesondere im Rahmen einer **Feldphase** (soweit erforderlich in mehreren Ländern) fort.

Bewährt haben sich eine erste Datenanalyse vor Ort, deren Validierung im Zuge eines Workshops mit Wissensträgern (Botschaft, lokale Vertretungen von Mittler- und Trägerorganisationen, Partnerorganisationen, ggf. andere Geber) und die zusammenfassende Vorstellung erster Ergebnisse ggf. mit Hinweisen auf erste Schlussfolgerungen (Präsentation).

Zum Abschluss der Feldmission(en) oder unmittelbar nach Rückkehr erstellt das Evaluierungsteam eine → „**Debriefing Note**“, d. h. einen formlosen Kurzbericht (z. B. per E-Mail) zu den in der Regel vor Ort vorgestellten vorläufigen Ergebnissen und ggf. Schlussfolgerungen (vgl. Präsentation zuvor) und dem erhaltenen Feedback (zumal zu gravierenden, abweichenden Sichtweisen), und übersendet diese der SG.

4.2.3 Fallstudien: Bei projektübergreifenden Evaluierungen wird in aller Regel nicht zu allen einschlägigen Projekten und Programmen eine Feldstudie als vertiefender Erhebungs- und Analyseschritt notwendig oder möglich sein. Vielmehr trifft das Evaluierungsteam eine kriterienbasierte Fallstudienauswahl (vgl. IR). Bei zusätzlich vorgesehenen, ressortspezifischen Fallstudien können zusätzliche Auswahlkriterien erforderlich sein.

Bei einer größeren Anzahl von Fallstudien, insbesondere wenn sie mit Feldstudien in mehreren Ländern verbunden sind, kann eine Arbeitsteilung im Team sinnvoll sein. Diese bedarf einer besonders sorgfältigen Vorbereitung und internen Qualitätssicherung durch die Teamleitung. Als unentbehrlich hat sich vielfach eine personengebundene „Wissensbrücke“ erwiesen. Bei einer kooperativen Durchführung empfiehlt sich, dass bei arbeitsteilig durchgeführten Primärerhebungen immer beide Projektpartner im Team vertreten sind.

Sofern mehrere Fallstudien im gleichen Land durchgeführt werden, muss eine Duplizierung von Evaluierungsmissionen vermieden werden, d. h., sie finden gleichzeitig und koordiniert statt und in der Vorbereitung und Durchführung dürfen Partnerorganisationen, andere Geber oder Zielgruppen nicht parallel, sondern nur arbeitsteilig oder gemeinsam kontaktiert, interviewt oder weitergehend beteiligt werden.

¹¹ Dazu sind ggf. vorgeschaltete Workshops erforderlich.

4.3 DATENANALYSE UND BERICHTSLEGUNG

Produkte: ggf. Fallstudienberichte; zusammenfassende Ergebnisse, Schlussfolgerungen und Empfehlungen (Bericht oder Präsentation) und Abschlussbericht (Entwurf, schließlich final); Stellungnahme der Ressorts

Austausch: ggf. RG-Treffen zu Fallstudienberichten (Entwurf) und zum Abschlussbericht

Das Evaluierungsteam setzt seine Datenanalyse nach Rückkehr von Feldmissionen fort.

4.3.1 Analyse von Fallstudien: Soweit mehrere Fallstudien zumal in mehreren Ländern durchgeführt werden, kann dazu, je nach Zielsetzung und Gegenstand, eine eigenständige Berichtslegung sinnvoll bzw. erforderlich sein. Soweit die Fallstudienenergebnisse nicht voll im Abschlussbericht integriert sind, empfehlen sich → **Fallstudienberichte**. Diese können entweder als Anlagen zum Abschlussbericht aufgenommen werden oder freistehende Berichte sein. Im ersten Fall enthalten sie keine Schlussfolgerungen und Empfehlungen. In besonderen Fällen, z. B. bei sehr komplexen Evaluierungsgegenständen, sind ggf. eigenständige Fallstudienberichte erforderlich.¹² Sofern neben ressortgemeinsamen Aspekten auch umfangreiche bzw. detaillierte, ressortspezifische Erkenntnisse und Empfehlungen angestrebt werden, die unterschiedliche Berichtswege nahelegen, können eigenständige, auf das jeweilige Ressort bezogene Fallstudienberichte ebenfalls sinnvoll sein. Schlussfolgerungen und insbesondere Empfehlungen können, sofern sie aufgenommen werden, dann nur sehr spezifisch sein und sollten erst nach oder gleichzeitig mit der Erstellung des Syntheseberichts finalisiert werden. Zu gewichtigen Fallstudien (Entwürfen) erfolgen interministerielle Referenzgruppentreffen; im Falle von ressortspezifischen Fallstudien und Berichten auf Wunsch zusätzlich auch getrennt.

Im Falle eines sich lediglich auf das BMZ-Portfolio beziehenden Fallstudienberichts gilt im Weiteren der übliche Ablauf von DEval-Evaluierungen, insbesondere Berichtsveröffentlichung, ggf. mit Anpassungen im Detail.

4.3.2 Vorbereitung und Finalisierung des Abschlussberichts: Bei ressortgemeinsamen Evaluierungen und damit dem Abschlussbericht steht das ressortübergreifende Erkenntnisinteresse an einem gemeinsamen Handlungsfeld im Vordergrund.

In anderen Zusammenhängen hat sich als Zwischenschritt vor der Ausformulierung des Abschlussberichts ein Kurzbericht oder eine Präsentation bewährt mit Blick auf → **Zusammenfassende Ergebnisse, erste Schlussfolgerungen und Empfehlungen** (Bericht oder Präsentation): Dieser Zwischenschritt dient zum einen der fachlichen Validierung der Ergebnisse; diese sollten daher eher ausführlich dargestellt¹³ und im ersten Teil eines RG-Treffens diskutiert werden. Zum anderen, in einem zweiten Teil, der Präsentation und Diskussion der daraus resultierenden Schlussfolgerungen und Empfehlungen, letztere ggf. auch in Form von Optionen.

Im Anschluss an das Referenzgruppentreffen und ggf. weiterer (gebündelter) Stellungnahmen erstellt das Evaluierungsteam den → **Abschlussbericht**: Evaluierungsberichte stellen Analysen und Bewertungen unabhängiger Expertenteams dar und entsprechen daher nicht notwendigerweise der Meinung oder eigenen Bewertungen der Ressorts. Sie orientieren sich an den international üblichen Qualitätsstandards. Insbesondere müssen die Evaluierungsfragen vollständig beantwortet und die Schlussfolgerungen und Empfehlungen logisch aus den Ergebnissen abgeleitet sein. Empfehlungen enthalten eindeutige Adressaten (z. B. „Ressorts“, ein bestimmtes, zu benennendes Ressort, Durchführungs- oder Förderorganisationen) und trennen möglichst strategische und eher operationelle Empfehlungen. Die Empfehlungen sollten umsetzbar und die Anzahl handhabbar (im Sinne von nachverfolgbar) sein.

Schutzwürdige Rechte von Betroffenen werden bei allen Berichtsfassungen berücksichtigt.¹⁴ Informationsquellen werden grundsätzlich, soweit nicht öffentlich zugänglich, anonymisiert, d. h., die Ausführungen dürfen nicht auf einzelne Interviewpartner oder vertrauliche Dokumente zurückführbar sein.

¹² Vgl. z. B. die internationale Gemeinschaftsevaluierung der Paris Declaration oder die Produkte der Tsunami Evaluation Coalition (TEC). Auch die ressortgemeinsame Evaluierung der humanitären Hilfe im Ausland (2011) umfasste mehrere Fallstudienberichte.

¹³ Alternativ, wiewohl unüblich, wäre ein Berichtsentwurf *nur zu den Ergebnissen* denkbar.

¹⁴ Dabei werden insbesondere auch schutzwürdige Belange von Personen im Partnerland berücksichtigt.

Der Entwurf des Abschlussberichts wird in einem (letzten) Referenzgruppentreffen diskutiert und von der RG, gebündelt nach Ressort, auch schriftlich innerhalb eines angemessenen Zeitrahmens mit Blick auf die sachliche Korrektheit kommentiert. Die SG kommentiert, ggf. mit externer Unterstützung, mit Blick auf die Einhaltung evaluatorischer Qualitätsstandards.

Mit Übersendung des finalen Berichts dokumentiert das Evaluierungsteam den Umgang mit den erhaltenen Kommentaren anhand eines Antwortrasters.

4.3.3 Berichtsweg und Follow-up I: Der oder die Verantwortliche für den Abschlussbericht übersendet der SG den finalen Bericht, ggf. nach Formatvorgaben der Ressorts. Die Evaluierungsstellen legen den Bericht mit einer eigenen (in der SG abgestimmten) Bewertung ihren jeweiligen Hausleitungen vor, zusammen mit einer eigenständigen (von den zuständigen Abteilungsleitungen gebilligten) und veröffentlichungsfähigen → **gemeinsamen Stellungnahme** der thematisch federführenden Referate. Die Stellungnahme behandelt die Relevanz und Nützlichkeit der Ergebnisse und skizziert die geplante Umsetzung der Hauptempfehlungen, ggf. auch, welche nicht geteilt werden. Die SG hat hierbei eine qualitätssichernde Funktion mit Blick auf die korrekte Darstellung von Evaluierungsergebnissen und Empfehlungen. Die Billigung der jeweiligen Hausleitung bezieht sich auf die Stellungnahme der Ressorts und nicht auf die Bewertungen des Evaluierungsberichts.

4.4 QUALITÄTSSICHERUNG ÜBER ALLE PHASEN HINWEG

Die *Evaluierungsteamleitung* ist für die Qualität der Produkte und der damit verbundenen Prozesse (Datenerhebung und -auswertung) sowie die Einhaltung der Zeitplanung verantwortlich. Dabei hat sich eine firmen- bzw. organisationsinterne Qualitätssicherung bewährt. Die *Steuerungsgruppe* ist verantwortlich für die Prozessqualität mit Blick auf die Einbeziehung relevanter Betroffener zu den genannten Meilensteinen, beginnend mit dem Auftakttreffen bis hin zur Umsetzungsplanung (vgl. 4.5). Darüber hinaus erfüllt sie eine Qualitätssicherungsfunktion bei Produkten, insbesondere dem Abschlussbericht durch Kommentierung. Soweit grundlegende Anforderungen an Vollständigkeit und Güte (unter Bezug auf die DeGEval- und OECD DAC-Qualitätsstandards, soweit einschlägig) in ausreichendem Ausmaß erfüllt sind, werden die Berichte (Abschlussbericht, ggf. Fallstudienberichte) von der SG akzeptiert. Die Unabhängigkeit der Bewertungen bleibt dabei gewahrt. Die *Mitglieder der Referenzgruppe* liefern einen Beitrag durch Prüfung und Kommentierung der sachlichen Korrektheit der Darstellung von Sachverhalten.

Insbesondere bei Evaluierungen größerer (sachlicher oder politischer) Reichweite, wie in aller Regel bei ressortgemeinsamen Evaluierungen, empfiehlt sich eine zusätzliche, *externe Qualitätssicherung* aus inhaltlich-fachlicher und/oder aus methodisch-evaluatorischer Sicht. Dazu werden ein oder mehrere Externe mit entsprechend ausgewiesener Qualifikation als Review Panel von der SG, vorzugsweise auch anlässlich Referenzgruppentreffen, mit einbezogen, die die Evaluierungsprodukte schriftlich und/oder mündlich kommentieren sowie ggf. bei strittigen Fragen zu evaluatorischen Prinzipien und Standards beratend – bis hin zu vermittelnd – von allen Beteiligten hinzugezogen werden können.

4.5 TRANSPARENZ, VERBREITUNG UND FOLLOW-UP II

Produkte: ggf. Kurzfassung des Evaluierungsberichts; Umsetzungsplan

Austausch: zum Umsetzungsplan

4.5.1 Veröffentlichung und Verbreitung: Evaluierungsberichte werden, nachdem sie den Ressortleitungen zur Kenntnis gebracht wurden, *in toto* auch den externen am Prozess Beteiligten zugänglich gemacht und vorzugsweise auch vollständig publiziert. Sofern schützenswerte Belange dagegensprechen, wird zumindest eine → **Kurzfassung**, d. h. aussagekräftige Zusammenfassung in deutscher Sprache, die vom Evaluierungsteam eigenverantwortlich erstellt wird, auf den Webseiten der Ressorts veröffentlicht oder verlinkt und in Fällen relevanter Partnerbeteiligung den Partnern in der Hauptlandessprache zur Verfügung gestellt.

Der publizierte Bericht enthält, klar abgegrenzt, auch die Stellungnahme der Ressorts, soweit diese nicht getrennt zugänglich gemacht wird.

Die beteiligten Ressorts unterstützen die interne und externe Verbreitung der Evaluierungsergebnisse und -empfehlungen im Sinne des Wissenstransfers. Insbesondere soweit eine Evaluierung auf einen Wunsch aus dem Deutschen Bundestag zurückgeht, übersenden die Ressorts den Abschlussbericht dem jeweils zuständigen Ausschuss vor Veröffentlichung.

4.5.2 Umsetzung von Empfehlungen: Die für den Evaluierungsgegenstand jeweils thematisch federführenden Referate der Ressorts sind nach Abschluss einer Evaluierung für die Umsetzung der Empfehlungen verantwortlich.

Sie erstellen dazu, möglichst gemeinsam oder mit einem gemeinsamen Teil, eine → **Umsetzungsplanung** zu den Empfehlungen, ggf. unter Beteiligung weiterer Referate gemäß Geschäftsverteilungsplan sowie mit Unterstützung der SG bzw. der Evaluierungsstellen; optional kann auch das Evaluierungsteam in beratender Funktion hinzugezogen werden. In dem i. d. R. ressortinternen Umsetzungsplan wird dokumentiert, inwieweit die Empfehlungen geteilt werden (vollständig, tlw., nicht – bei tlw. oder nicht: mit Begründung) und welche Schritte zur Umsetzung der Empfehlungen geplant sind. Die Umsetzungsschritte sollten „SMART“ formuliert sein.

Der Umsetzungsplan ist von den zuständigen Abteilungs- oder vertretungsweise Unterabteilungsleitungen zu billigen.

Die thematisch federführenden Referate sorgen anschließend für die Umsetzung der geplanten Schritte.

Der Umsetzungsplan bildet die Grundlage für ein späteres **Monitoring** der SG bzw. der jeweiligen Evaluierungsstellen, inwieweit die vereinbarten Umsetzungsschritte tatsächlich realisiert wurden. Die Ergebnisse werden der Leitung des jeweiligen Ressorts zur Kenntnis gebracht.

5 Literatur

ALNAP (2006): Evaluating Humanitarian Action Using the OECD DAC Criteria.

ALNAP (2016): Evaluation of Humanitarian Action Guide.

DeGEval (2016): Standards für Evaluation; 1. Revision.

EU (2017): Joint statement by the Council and the representatives of the governments of the Member States meeting within the Council, the European Parliament and the Commission „THE NEW EUROPEAN CONSENSUS ON DEVELOPMENT“ (2017/C 210/01).

OECD DAC (1991): Principles for Evaluation of Development Assistance.

OECD DAC (2006): Guidance for Managing Joint Evaluations.

OECD DAC (2010): Qualitätsstandards für die Entwicklungsevaluierung.

OECD DAC (2012): Evaluating Peacebuilding Activities in Settings of Conflict and Fragility.

OECD DAC (2019): DAC Recommendation on the Humanitarian-Development-Peace Nexus.

OECD DAC (2019): Better Criteria for Better Evaluation; Revised Evaluation Criteria Definitions and Principles for Use.

UNEG (2013): Resource Pack on Joint Evaluations.

World Humanitarian Summit (2016): Commitments to Action.