

Die Servicestelle Evidenz des BMZ

Ein Ausgestaltungsvorschlag auf Grundlage internationaler Erfahrungen

www.kooperationglobal.de



Executive Summary

Der BMZ-Reformplan *Zukunft zusammen global gestalten* vom Januar 2026 verankert Evidenz und Wirksamkeit als zentrale Steuerungsprinzipien der deutschen Entwicklungszusammenarbeit. Die darin angekündigte Servicestelle Evidenz bietet eine vielversprechende Möglichkeit, die Entscheidungsgrundlagen des BMZ umfassend zu stärken. Die **Wirksamkeit** der vom BMZ gesteuerten deutschen EZ-Mittel – für 2027 etwa 9,5 Milliarden Euro – könnte durch sie **signifikant gesteigert** werden.

Das vorliegende Analysepapier legt auf Grundlage internationaler Erfahrungen konkrete Ausgestaltungsvorschläge für die Servicestelle Evidenz vor. Es richtet sich an die Verantwortlichen im BMZ, die mit der Konzeption und dem Aufbau der Servicestelle betraut sind, an Kooperationspartner der Fachlichkeiten Evidenz und Wirksamkeit in GIZ, KfW und DEval als zentralen Partnerinstitutionen, und an die interessierte Öffentlichkeit. Die Analyse beruht auf einer vergleichenden Untersuchung von **elf internationalen Evidenzeinheiten bzw. -programmen** in sechs Ländern und zwei multilateralen bzw. internationalen Organisationen, die wie folgt zusammengefasst werden kann:

Internationale Beispiele – Lehren aus der Praxis

US: USAID Office of the Chief Economist (OCE): Rd. 1,7 Mrd. USD in wirksamere Programme umgeschichtet durch systematische Kosten-Effektivitäts-Analysen. Team von über 30 Mitarbeitenden. Zeigt, dass eine Evidenzeinheit mit starkem Mandat und Zugang zur Leitungsebene enorme Hebelwirkung entfalten kann.

UK: FCDO Analysis Directorate und angeschlossene Evidenzstrukturen: 15-jährige Erfahrung und Praxis mit Kosten-Nutzen-Rechnungen für EZ-Projekte, vertiefende Evidenzchecks für Projektvorschläge ab bestimmten Schwellenwerten, dazu drei weitere Arbeitsstränge zu Evidenz auf operativer, strategischer und evaluativer Ebene.

CH: Ressort Qualität und Ressourcen des SECO-WE: Seit 2025 Pilotprojekt für einen Rapid-Evidence-Review Service; Ausarbeitung von Leitlinien zur Anwendung von (kausalen) Evidenz- und Evaluierungsmethodiken, die Programmentscheidungen stärker beeinflussen, infolge einer Metaevaluierung zur Wirksamkeit im eigenen Portfolio.

NOR: Development Learning Lab (DLL): Praxisnahe Zusammenarbeit zwischen Forschung und Durchführungsorganisationen. Besonders relevant für das evidenzorientierte Zusammenspiel BMZ–GIZ–KfW und den gezielten Transfer von Erkenntnissen aus der Forschung. Zusätzlich Schnittstelle zur Einbindung zivilgesellschaftlicher Akteure in der Evidenzarchitektur.

UGA & ZAF: Center for Rapid Evidence Synthesis (ACRES) und Pan-African Collective for Evidence (PACE): Evidenzinfrastrukturen im Globalen Süden, die zeigen, wie KI-gestützte Rapid-Evidence-Services und Kapazitätsaufbau in Partnerländern aussehen und mit Evidenzinfrastruktur aus Geberländern zusammenarbeiten können.

Multilateral: Development Impact Evaluation (DIME) Unit und Strategic Impact Evaluation Fund (SIEF) der Weltbank: Gezielte und rigorose Wirkungsevaluierungen (u.a. RCTs) in Ländern mit geringem und mittlerem Einkommen, Zusammenführen der Ergebnisse in einer Plattform und Bereitstellung von KI-Lösungen zu evidenzbasiertem Arbeiten (ImpactAI).

International: Airbel Impact Lab (AIL) des International Rescue Committee (IRC): Liefert weit überproportionalen Anteil an Wirkungsevaluierungen im humanitären Sektor und fokussiert auf Kosteneffizienz- und Kosteneffektivitätsanalysen. Entwicklung mehrerer web-basierter Tools zur Einbindung der Ergebnisse in operative Bereiche.

Die **zentralste Erkenntnis aus der internationalen Praxis**: Evidenzeinheiten entfalten nur dann Wirkung, wenn sie die zentralen Steuerungsentscheidungen vor der Programmkonzeption beeinflussen können. Werden sie lediglich nachgelagert oder rein beratend eingesetzt, bleibt ihr Einfluss auf eine wirkungsorientierte Programmsteuerung und Mittelallokation marginal.

Entsprechend leitet dieses Analysepapier für die Ausgestaltung der Servicestelle Evidenz des BMZ **drei Kernempfehlungen** ab:

1. Strategische Verankerung an der Leitungsebene: Die Servicestelle sollte direkt an die Abteilung für Wirksamkeit & Daten (Abteilung G) angebunden sein und mindestens einen Stabsrang mit direkter Berichtslinie an die Leitungsebene haben, um sicherzustellen, dass Evidenz in zentrale Steuerungsentscheidungen einfließt. Bestenfalls enthält sie auch die Rolle eines „Leitenden Wirkungsanalysten“ (LWA), der sie stark nach innen und außen vertritt und im Auftrag der Leitungsebene den erneuerten Wirksamkeitsfokus des BMZ verkörpert.

2. Klare Mandate und Funktionen mit messbaren Ergebnissen: Die Servicestelle sollte mit verschiedenen Kernfunktionen die evidenz- und wirksamkeitsorientierte Arbeit des BMZ stärken, u.a. mit (1) Analysen und Voten zur wirkungsorientierten Haushaltssteuerung, (2) einem verbindlichen Evidenz- und Kosten-Effektivitäts-Check bei Programmkonzeptionen mit hohen Mittelvolumina, (3) einem regelmäßigen Portfolio-Wirksamkeitsbericht und (4) einem Rapid-Evidence-Helpdesk.

3. Angemessene Ressourcenausstattung: Die Servicestelle Evidenz benötigt zur Erfüllung ihrer Aufgaben kurzfristig mindestens 7 Vollzeitäquivalente und eine Mittelausstattung für Programme strategischer Wirkungsevaluierungen, sowie zur Entwicklung von (internen) Informations- und Kollaborationsplattformen. Flankiert werden sollte dies durch eine Erhöhung des Titels für Forschung und Evaluierung im Einzelplan 23. Eine solche Ausstattung entspräche etwa 0,5% des gesamten BMZ-Haushalts, die sich durch Hebelungseffekte einer stärkeren Evidenz, Kosteneffektivität und Wirksamkeit im Gesamtportfolio des BMZ schnell um ein Vielfaches amortisieren würden.

Die Servicestelle ist ein geeignetes **Vehikel des organisationalen und gesellschaftlichen Lernens** zu Evidenz und Wirksamkeit im BMZ. Sie kann im Ressort den Kapazitätsaufbau und die Schaffung einer Lernkultur unterstützen und durch strategische Wirkungsevaluierungen genau dort robustes Wissen generieren, wo das Ministerium im komplexen Feld der Entwicklungszusammenarbeit noch Evidenzlücken sieht.

Die vielen internationalen Beispiele und Ambitionen von Geberländern und Partnern zeigen: Deutschland ist bei diesen Bestrebungen nicht allein. Die Servicestelle sollte gezielt zu Evidenz und Wirksamkeit in den **internationalen Austausch** treten, zum Beispiel über Netzwerke wie die Smart Buys Alliance.

Nicht vergessen werden sollte auch, dass es **in der deutschen EZ bereits vielversprechende Praxisbeispiele** einer evidenzbasierten Steuerung gibt, auf denen aufgebaut werden kann, und auf die in diesem Analysepapier ebenfalls in einem kurzen Exkurs geblickt wird.

In Summe bestätigt sich das Potential der Servicestelle Evidenz als einer der zentralen Bausteine zur Realisierung der im BMZ-Reformplan formulierten Ansprüche und Pläne für eine zukunftsfähige deutsche Entwicklungszusammenarbeit.

Inhaltsverzeichnis

Executive Summary.....	2
1. Ausgangslage und Reformkontext.....	4
2. Internationale Vorbilder: Evidenzeinheiten in der Entwicklungszusammenarbeit.....	6
2.1 USAID Office of the Chief Economist (USA).....	6
2.2 FCDO Analysis Directorate und angeschlossene Evidenzstrukturen (Vereinigtes Königreich).....	8
2.2.1 K4DD – Knowledge for Development and Diplomacy.....	9
2.2.2 SIEL – Strategic Impact Evaluation and Learning.....	10
2.2.3 ICAI – Independent Commission for Aid Impact.....	11
2.3 WEQA – Einheit Qualität und Ressourcen, SECO-WE (Schweiz).....	12
2.4 DLL – Development Learning Lab (Norwegen).....	13
2.5 ACRES – Center for Rapid Evidence Synthesis (Uganda).....	15
2.6 PACE – Pan-African Collective for Evidence (Südafrika).....	16
2.7 DIME und SIEF – Development Impact Evaluation und Strategic Impact Evaluation Fund (Weltbank).....	16
2.8 AIL - Airbel Impact Lab des International Rescue Committee (IRC).....	18
2.9 Vergleichende Übersicht.....	19
3. Empfohlene Mandate und Funktionen der Servicestelle.....	20
3.1 Funktion 1: Wirkungsorientierte Haushaltssteuerung.....	21
3.2 Funktion 2: Verbindlicher Evidenz- und Kosten-Effektivitäts-Check.....	21
3.3 Funktion 3: Konzeptionelle Arbeiten und Portfoliobericht zu Evidenz- und Wirksamkeit im BMZ.....	22
3.4 Funktion 4: Rapid-Evidence-Helpdesk.....	23
3.5 Funktion 5: Kapazitätsaufbau und Lernkultur.....	24
3.6 Funktion 6: Strategische Beauftragung von Wirkungsevaluierungen.....	24
3.7 Funktion 7: Vernetzung und internationaler Austausch.....	25
3.8 Funktion 8: Aufbau einer Wirksamkeitsdatenbank für die dt. EZ.....	25
3.9 Funktion 9: Strategische Kommunikation von Wirksamkeit.....	25
Exkurs: Evidenzbasierte Steuerung in der Praxis der deutschen EZ.....	26
4. Institutionelle Ausgestaltung.....	28
4.1 Klare Funktionsabgrenzung.....	28
4.2 Institutionelle Verankerung.....	28
4.3 Personal und Ressourcenausstattung.....	30
4.3 Erfolgsfaktoren und Risiken.....	31
4.3.1 Erfolgsfaktoren.....	31
4.3.2 Risiken und Gegenmaßnahmen.....	31
5. Schlussfolgerung.....	32
Quellenverzeichnis.....	33
Annex: Umsetzungszeitplan.....	36

1. Ausgangslage und Reformkontext

Die Mittel des BMZ wurden in den vergangenen fünf Jahren um fast ein Drittel gekürzt, während die globalen Herausforderungen dramatisch gewachsen sind. Damit wird die effiziente und effektive Verwendung der verbleibenden Mittel noch wichtiger. Der BMZ-Reformplan vom Januar 2026 reagiert auf diese Spannung mit einem klaren Bekenntnis zu evidenzbasierter Steuerung. Ministerin Alabali Radovan kündigte an, Investitionen in evidenzbasierte Ansätze zu verstärken und insgesamt stärker nach Wirkung zu steuern.

Der Reformplan verspricht, wissenschaftliche Evidenz systematischer zu nutzen und „weniger wirksame Vorgehensweisen zugunsten evidenzgestützter Alternativen schrittweise zurückzuführen.“ Er setzt damit ein starkes Zeichen: Das BMZ wird Programme auf Grundlage vergleichender Wirksamkeitsanalysen umsteuern oder beenden – und reagiert damit auf einen Wirksamkeitsanspruch, den auch die deutsche Bevölkerung ganz besonders hervorhebt.¹

Die angekündigte Servicestelle Evidenz ist das zentrale institutionelle Instrument, um diesen Anspruch einzulösen. Im Reformplan wird sie als Teil einer neuen Architektur beschrieben, die Evidenz als Grundlage von Allokationsentscheidungen, zur Analyse von Länderportfolien und für die Ergebnisberichterstattung stärker etabliert. Des Weiteren sollen GIZ und KfW beauftragt werden, Standardwirkungsmodelle einzuführen, und übergreifend Datenverfügbarkeit, -management und -qualität ausgebaut werden. Die Umsetzung der Reform soll bis Sommer 2027 abgeschlossen sein.²

Der Reformplan verweist zudem auf das bereits existierende Wissensökosystem, welches das BMZ umgibt: DEval, GIGA, IDOS, sowie weltweite Forschungspartnerschaften und Kooperationen mit Universitäten und Think Tanks. Die Servicestelle sollte dieses Ökosystem nicht ersetzen, sondern als zentraler Knotenpunkt dafür sorgen, dass es gepflegt und wo nötig strategisch erweitert wird, und dass die vorhandene Evidenz systematisch in Entscheidungsprozesse des BMZ einfließt. Dieses Papier analysiert, wie andere Geberländer, Organisationen und Partner genau diesen Schritt bereits vollzogen haben, und leitet daraus konkrete Empfehlungen für das BMZ ab.

¹ DEval (2026). Meinungsmonitor Entwicklungspolitik. [\[Link\]](#)

² BMZ (2026). Zukunft zusammen global gestalten. S. 24f. [\[Link\]](#)

2. Internationale Vorbilder: Evidenzeinheiten in der Entwicklungszusammenarbeit

Die folgenden elf internationalen Modelle bieten konkrete Lehren für die Ausgestaltung der Servicestelle Evidenz in Deutschland. Sie decken das gesamte Spektrum relevanter Funktionen ab: Von strategischer Haushaltssteuerung und Kosten-Effektivitäts-Analysen³ über Rapid-Evidence-Services und rigorosen Wirkungsevaluierungen bis hin zu praxisnaher Zusammenarbeit mit Durchführungsorganisationen und der partnerschaftlichen Koordination mit Evidenzinfrastruktur im Globalen Süden.

Für jede Institution wird dargelegt, welche Elemente für den deutschen Kontext besonders relevant und übertragbar sind – stets unter Berücksichtigung der spezifischen Strukturen der deutschen EZ mit ihren Durchführungsorganisationen, dem Zusammenspiel von bilateraler und multilateraler Zusammenarbeit und den bestehenden institutionellen Partnerschaften.

Eine Gesamtübersicht aller Beispiele in kompakter tabellarischer Form findet sich in [Unterkapitel 2.9](#).

2.1 USAID Office of the Chief Economist (USA)

Das im Jahr 2023 unter der Leitung von Prof. Dean Karlan (Northwestern University) eingerichtete, unabhängige Office of the Chief Economist (OCE) war die erste zentrale Evidenzeinheit innerhalb von USAID. Mit einem Team von über 30 Mitarbeitenden⁴ hatte das OCE den Auftrag, Kosteneffektivitätsanalysen und beste verfügbare Evidenzen systematisch in die Entscheidungsfindung und Programmgestaltung von USAID einzubinden.⁵ Das OCE konnte auf bestehenden Bemühungen aufbauen: Bereits 2021 hatte USAID in seinen Verwaltungsrichtlinien verankert, dass Kostenanalysen Teil aller USAID-finanzierten Wirkungsevaluierungen sein müssen.⁶ Zusätzliches Kernziel des OCE war, neue Evidenz zu kosteneffektiven Interventionen in Handlungsfeldern mit noch bestehenden Evidenzlücken zu generieren.⁷

Innerhalb von rund zweieinhalb Jahren gelang es dem OCE laut eigener Aussage, ungefähr 1,7 Milliarden USD in Programme mit stärkerer Evidenzbasis umzuschichten.⁸ Um als zunächst kleines Team und mit wenigen Ressourcen dennoch schnell handlungsfähig zu sein, operierte das OCE als eine Art flexible und nachfrageorientierte Beratungseinheit zu

³ Kooperation Global unterscheidet in diesem Analysepapier zwischen Kosteneffizienz und Kosteneffektivität. Ersteres meint „die Dinge richtig zu tun“, d.h. den Ressourceneinsatz für gleichbleibende Outputs zu optimieren. Kosteneffektivität hingegen zielt ab darauf „die richtigen Dinge zu tun“, also vor allem die Maßnahmen mit dem besten Verhältnis von Kosten und *Wirkungen* zu identifizieren. Die Servicestelle Evidenz sollte beide Konzepte in ihren Analysen abdecken. Da sie aber nicht selbst implementiert, sollte sie ein besonderes Augenmerk auf die Kosteneffektivität legen.

⁴ Zur besseren Lesbarkeit wird in diesem Analysepapier auf eine geschlechtsspezifische Differenzierung verzichtet. Alle personenbezogenen Bezeichnungen gelten gleichermaßen für alle Geschlechter.

⁵ USAID (2024). USAID Position Paper: Cost-Effectiveness. [\[Link\]](#)

⁶ USAID (2021). ADS Chapter 201. Program Cycle Operational Policy. [\[Link\]](#)

⁷ Vgl. USAID (2024). Promoting Impact and Learning with Cost-Effectiveness Evidence (PILCEE). [\[Link\]](#)

⁸ Ruiz (2025). How to fix foreign aid - Interview with Dean Karlan. [\[Link\]](#)

Evidenz („Evidence Use Teams“) innerhalb von USAID.⁹ Die Priorisierung bestimmter Handlungsfelder und Programme zur vertieften Arbeit mit Evidenz und Wirksamkeit erfolgte unter drei Kriterien: Sie mussten eine gewisse Größe im Bezug auf das Mittelvolumen haben, es musste noch eine relevante Lücke zwischen aktueller Programmgestaltung und verfügbarer Evidenz bestehen, und die Fachreferate und Durchführungsorganisationen mussten selbst motiviert zur Kollaboration sein.¹⁰

Den eingeleiteten Paradigmenwechsel beschreibt das OCE so: USAID habe historisch Programme vor allem anhand von Rechenschaftskriterien evaluiert und gesteuert – also etwa, ob die vorgesehene Zahl an Begünstigten erreicht wurde. Die eigentliche Frage müsse jedoch lauten, ob das Programm tatsächlich kurz-, mittel- und langfristig Wirkungen bei den Menschen erzielt habe und die besten verfügbaren Evidenzen und Ansätze nutzte.¹¹ Dieser Unterschied zwischen einer prozessorientierten und einer wirkungsorientierten Evaluierung und Steuerung ist auch für das BMZ zentral.

Nach der Abwicklung von USAID im Jahr 2025 gründeten ehemalige Mitarbeitende des OCE die Project Resource Optimization (PRO) Initiative, die zunächst beim Center for Global Development angesiedelt war und nun in den DIV-Fund überführt wurde.¹² PRO zeigt, wie die im OCE aufgebauten Kompetenzen – insbesondere die systematische Identifikation besonders kosteneffektiver Programme – auch außerhalb staatlicher Strukturen Wirkung entfalten können. Innerhalb weniger Monate mobilisierte PRO über 125 Millionen USD an philanthropischem Privatkapital für besonders wirksame Programme.

Auch die aktuelle US-Administration betont für ihre Neugestaltung der amerikanischen Entwicklungszusammenarbeit die Relevanz von Kosteneffizienz und Transparenz. Im verabschiedeten Haushaltsgesetz 2026 wird das State Department ausdrücklich dazu aufgefordert, Strategien zur Verbesserung von Effektivität und zur Einbeziehung vergangener Wirkungsevaluierungen in der amerikanischen EZ vorzulegen.¹³ Trotz einem institutionellen Bruch durch die Auflösung von USAID ist daher eine Kontinuität von Ansätzen des OCE nicht unwahrscheinlich.

Relevanz für das BMZ: Das OCE-Modell zeigt, welche Hebelwirkung eine Evidenzeinheit mit starkem Mandat und direktem Zugang zur Leitungsebene entfalten kann. Die Servicestelle sollte ähnlich angelegt sein: Nicht als rein akademische Beratung, sondern als operatives Instrument der Mittelvergabe. Besonders drei OCE-Praktiken sind direkt übertragbar: (1) der verbindliche Kosten-Effektivitäts-Check als Standardschritt in der Programmkonzeption, (2) die Einrichtung eines Evidence Use Teams, das vorhandene Evidenz systematisch für laufende Programmkonzeptionen aufbereitet (in der deutschen EZ besonders durch Durchführungsorganisationen zu erfüllen), und (3) die direkte Berichterstattung an die Leitungsebene.

⁹ Karlan, D., Hanney, O. (2026). Dean Karlan on USAID, evidence-based policy, and key research gaps. [\[Link\]](#)

¹⁰ Ibid.

¹¹ USAID (2024). Promoting Impact and Learning with Cost-Effectiveness Evidence (PILCEE). [\[Link\]](#)

¹² CGD (o. J.). Project Resource Optimization (PRO). [\[Link\]](#)

¹³ US Kongress (2026). H.R.7148 - Consolidated Appropriations Act, 2026. Sec. 7011. [\[Link\]](#)

2.2 FCDO Analysis Directorate und angeschlossene Evidenzstrukturen (Vereinigtes Königreich)

Das Analysis Directorate des Foreign, Commonwealth and Development Office (FCDO) des Vereinigten Königreichs wurde 2020 aufgesetzt und ist eine Einheit mit über 100 Mitarbeitenden, die insbesondere auch analytische Fachprofile haben (Datenanalyse, Verhaltensforschung, Evaluierung, Ökonometrie, Statistik, OSINT¹⁴, u.a.). Ihm steht ebenfalls die Rolle eines Chief Economists vor, die aktuell von Dennis Novy (University of Warwick) bekleidet wird.¹⁵ Erklärtes Hauptziel des Analysis Directorate ist es, den Qualitätsstandard des FCDO hinsichtlich ökonomischer Analyse, Fachplanung und Evidenznutzung für die Programm- und Projektplanung zu heben.¹⁶

Bereits 2011 begann das ehemalige Department for International Development (DFID) mit strukturierten Kosten-Nutzen-Analysen (Value for Money, VfM) für die EZ-Programme des Vereinigten Königreichs.¹⁷ Es war damit sowohl unter den Ministerien des Vereinigten Königreichs, als auch weltweit einer der Vorreiter.¹⁸ Im Sinne verstärkter Evidenz- und Wirksamkeitsanalysen finanzierte das DFID seitdem etwa 30% aller randomisierten Kontrollstudien (RCT) mit einem entwicklungsökonomischen Fokus, die überwiegend in den besten fünf ökonomischen Fachzeitschriften der Welt veröffentlicht wurden.¹⁹

In den vergangenen Jahren wurden diese Analysen und Praktiken im FCDO schrittweise ausgeweitet und in sogenannten „Best Buys“ und „Great Buys“, also Programmen und Projekten mit herausragender Kosteneffektivität, zusammengeführt.²⁰

Für Projektanträge mit einem Volumen von mehr als 40 Millionen Pfund wurde ein vertiefter Qualitäts- und Kosteneffektivitätscheck eingeführt, in welchem Referenten zunehmend auch auf „Best Buys“ als evidenzbasierte Referenzen verweisen, oder aber gegen diesen klaren Vergleichsmaßstab begründen, warum sie sich für andere Projektgestaltungen aussprechen.²¹ Die detaillierten Leitlinien zur Programmierung der EZ durch das FCDO sind öffentlich einsehbar.²²

Das Analysis Directorate hat ein weitergehendes Aufgabenfeld als es für die Servicestelle Evidenz durch den BMZ-Reformplan angedeutet wird und ist mit seinen etwa 100 Mitarbeitenden nicht 1:1 auf die zukünftige Evidenzarchitektur des BMZ und seiner Durchführungsorganisationen übertragbar. Manche seiner Analyseaufgaben (und Personalressourcen) könnten im deutschen Fall auch eher von den Fach- und Methodenbereichen bzw. Plattformen von Fachlichkeiten in der GIZ oder KfW abgedeckt werden. Es ist jedoch hinsichtlich einer sukzessiven Ausweitung von Evidenz- und Kosteneffektivitätsanalysen in die Programm- und Arbeitsebene der EZ beispielgebend, und für FCDO die entscheidende Institutionalisierung der unmittelbaren Übersetzung von Evidenz und Wirksamkeitsanalysen in zukünftige Programmentscheidungen.

¹⁴ Open-Source Intelligence, d.h. die Informationsgewinnung aus offenen Quellen, siehe z.B.

¹⁵ Gov.uk (n.d.). Chief Economist and Director of Analysis Professor Dennis Novy. [\[Link\]](#)

¹⁶ FCDO (2024). Chief Economist and Director of Analysis - Candidate Brief. [\[Link\]](#)

¹⁷ DFID (2011). DFID's Approach to Value for Money (VfM). [\[Link\]](#)

¹⁸ International Development Committee of the UK Parliament (2025). Assessing Value, Ensuring Impact: The FCDO's Approach to Value for Money in Official Development Assistance. [\[Link\]](#)

¹⁹ Leight, J. (2022). Patterns of randomization, geographic concentration and funding in the recent development economics literature. [\[Link\]](#)

²⁰ Ritchie, E. (2025). FCDO's Best Buys Deserve a Larger Audience. [\[Link\]](#)

²¹ Ibid.

²² FCDO (2026). FCDO Programme Operating Framework. [\[Link\]](#)

Markant am Beispiel des Vereinigten Königreichs ist auch die Einbettung in einen „Whole-of-government-approach“, einen koordinierten ressortübergreifenden Ansatz: In der gemeinsamen öffentlichen Datenbank spezifischer Forschungsinteressen aller Regierungsbehörden des UK, der ARI-Database, hält FCDO aktuell etwa 180 Einträge.²³ Das in Deutschland neu gegründete Bundesministerium für Digitales und Staatsmodernisierung (BMDS) bietet die Möglichkeit, solche Ansätze zur *übergeordneten* Entwicklung von Evidenz und Wirkungsorientierung auch hierzulande voranzutreiben.

Spezifisch für die EZ koordiniert FCDO darüber hinaus drei weitere Säulen der Evidenzarchitektur des Vereinigten Königreichs, die ebenfalls für ein zukünftiges Modell in Deutschland wertvolles Anschauungsmaterial bieten:

2.2.1 K4DD – Knowledge for Development and Diplomacy

K4DD ist ein FCDO-finanziertes Programm, das seit April 2024 in seiner aktuellen Phase läuft (November 2023 bis Oktober 2027, Gesamtlaufzeit 3,5 Jahre).²⁴ Es wird von einem Konsortium aus sechs Organisationen unter Führung des Institute of Development Studies (IDS) getragen.

K4DD produziert jährlich über 200 Rapid-Evidence-Reviews für FCDO-Mitarbeitende und bietet darüber hinaus Lernformate und Wissensprodukte an. Der Service ist nachfrageorientiert: FCDO-Referenten können über ein Helpdesk-System schnelle Evidenzsynthesen zu spezifischen Fragen anfordern. Ein typischer Rapid-Evidence-Review umfasst hier bisher sechs Arbeitstage und wird innerhalb eines Monats geliefert. Für komplexere Fragen stehen sogenannte „Emerging Issues Reports“ zur Verfügung, die über einen längeren Zeitraum erarbeitet werden.

James Georgalakis, ehemaliger Leiter von K4DD, betont als zentrale Lehre aus über 25 Jahren Erfahrung (K4DD baut auf Vorgängerprogrammen wie K4D auf), dass der Prozess der Evidenznutzung ebenso wichtig sei wie die Qualität der Evidenz selbst.²⁵ Ein Drittel der Anfragen für Rapid-Evidence-Reviews betreffe nicht mehr ODA-Themen im engeren Sinne, sondern auch fachliche Fragen hinsichtlich internationaler Sicherheitspolitik, Wissenschaftsdiplomatie, Industriepolitik oder Migration.²⁶ Diese Komplexität müssen Evidenzeinheiten durch ein hohes Maß an Agilität und Flexibilität navigieren können. Neben klassischen Literaturanalysen sollten sie daher auch andere Methoden der Evidenzgenerierung in ihr Instrumentarium aufnehmen, etwa die Einbeziehung individueller und idiosynkratischer Expertise oder die Analyse unstrukturierter bzw. unkonventioneller Daten.

Relevanz für das BMZ: K4DD ist das beste internationale Modell für den Rapid-Evidence-Service, den die Servicestelle anbieten sollte. Für das BMZ könnte ein ähnliches Konsortium aus DEval, GIGA, IDOS und ausgewählten Universitäten und Instituten die Trägerschaft des Rapid-Evidence-Helpdesks übernehmen. Die K4DD-Erfahrung zeigt auch: Ein solcher Service sollte von Anfang an alle Berichte (soweit möglich) öffentlich zugänglich machen – dies schafft ein öffentliches Gut und erhöht die Sichtbarkeit des BMZ als evidenzbasierter Akteur.

²³ ARI-Database (n.d.). Areas of Research Interests and Questions of the FCDO. [\[Link\]](#)

²⁴ K4DD (2026). The Knowledge for Development and Diplomacy Programme. [\[Link\]](#)

²⁵ Georgalakis, J. (2026). Navigating Shifting Policy Priorities in Global Evidence Synthesis. [\[Link\]](#)

²⁶ Ibid.

2.2.2 SIEL – Strategic Impact Evaluation and Learning

SIEL ist ein sechsjähriges Programm (2024–2030), getragen von einer Partnerschaft zwischen der Evaluation Unit des FCDO, Innovations for Poverty Action (IPA) und dem Abdul Latif Jameel Poverty Action Lab (J-PAL).²⁷

SIEL finanziert rigorose Wirkungsevaluierungen – einschließlich randomisierter kontrollierter Studien (RCTs), quasi-experimenteller Methoden und Langzeit-Follow-up-Studien – in vier strategischen Prioritätsbereichen des FCDO: Wachstum, humanitäre Hilfe, Klima und Natur sowie Konflikt und Fragilität. SIEL hat für jeden dieser Bereiche Learning Agendas entwickelt, die zentrale Evidenzlücken identifizieren und künftige Evaluierungen gezielt steuern.

Das Budgetmodell von SIEL ist aufschlussreich: 60 Prozent fließen in Wirkungsevaluierungen, 15 Prozent in Langzeit-Follow-up-Studien und 15 Prozent in Kapazitätsaufbau.²⁸ FCDO-Mitarbeitende können sich über halbjährliche Ausschreibungsrunden um Evaluierungsfinanzierung bewerben. SIEL erzeugt nicht nur Evidenz, sondern investiert erheblich in deren Nutzung – durch maßgeschneiderte Schulungen, Disseminationsstrategien und direkte Zusammenarbeit mit Programnteams bei der Interpretation und Anwendung von Ergebnissen.

Auch das BMZ hat mit dem RIE-Förderprogramm im DEval²⁹ einen ersten ähnlichen Schritt in Richtung stärkerer rigoroser Wirkungsevaluierungen getan und dabei die Antragsberechtigung sogar weiter als SIEL gesetzt. Nach Ablauf des RIE- Förderprogramms im Juni 2026 sollten Nachfolgeprogramme aufgesetzt werden, die vom Evaluierungsreferat und der Servicestelle Evidenz strategisch koordiniert werden.

Die Architektur des UK-Systems gibt eine klare Rollenverteilung vor: Während K4DD vorhandene Evidenz für zeitnahe Entscheidungen aufbereitet (evidence synthesis), erzeugt SIEL neue Evidenz durch rigorose Evaluierungen (evidence generation). Beide Programme sind in der Evaluation Unit des FCDO verankert, die wiederum im Analysis Directorate unter Leitung des Chief Economist angesiedelt ist.

Relevanz für das BMZ: SIEL zeigt, wie eine Evidenzeinheit systematisch neue Wirkungsevidenz erzeugen kann – komplementär zur Aufbereitung vorhandener Evidenz (wie im K4DD-Modell). Für das BMZ könnte ein ähnliches Partnerschaftsmodell mit IPA, J-PAL oder vergleichbaren Organisationen die strategische Beauftragung von Wirkungsevaluierungen bündeln und an die Bedarfe der Servicestelle koppeln. Die Learning-Agenda-Methodik ist direkt übertragbar: Die Servicestelle könnte für jeden der vier Reformziele des BMZ Evidenzlücken identifizieren und entsprechende Evaluierungen priorisieren.

²⁷ Innovations for Poverty Action (2024). Strategic Impact Evaluation and Learning (SIEL). [\[Link\]](#)

²⁸ FCDO (2023). Strategic Impact Evaluation and Learning (SIEL) programme – Call for Proposals. [\[Link\]](#)

²⁹ DEval (o. J.). Das RIE-Förderprogramm. [\[Link\]](#)

2.2.3 ICAI – Independent Commission for Aid Impact

Die ICAI ist die unabhängige Evaluierungsinstitution für die britische Entwicklungszusammenarbeit. Ihre gesetzliche Grundlage ist der International Development Act 2015. Mit drei Beauftragten und rund 10 Mitarbeitenden führt sie thematische Reviews der ODA-Ausgaben durch und berichtet an das International Development Committee des Parlaments. Die Regierung ist verpflichtet, innerhalb von sechs Wochen auf ICAI-Berichte zu reagieren.³⁰

Die FCDO-Review der ICAI aus dem Jahr 2020 empfahl, die ICAI stärker als „kritischen Freund“ (critical friend) zu positionieren: Unabhängig und unvoreingenommen, mit dem Ziel, die Wirksamkeit der EZ zu verbessern, und nicht nur Teil der Rechenschaftspflicht zu sein. Ein zentraler Reformvorschlag war die Ernennung eines Senior Official im FCDO, der als Ansprechperson für die ICAI dient und sicherstellt, dass Empfehlungen tatsächlich umgesetzt werden.³¹

Funktional kommt ICAI einer Kombination aus DEval und Bundesrechnungshof am nächsten: Neben der fachlichen Evaluierung legt es einen starken Schwerpunkt auf Kosteneffizienz, zudem berichtet es direkt an das Parlament. Die letzteren beiden Funktionen würden im „deutschen Modell“ der Evidenzarchitektur durch die Servicestelle gestärkt werden.

In seiner stärker auf Parlament und Öffentlichkeit ausgerichteten Berichterstattung nutzt ICAI zudem ein klares Ampelsystem zur ex-post Evaluierung von Projekten³² und nimmt standardmäßig explizit lokale Stimmen aus den Zielländern von Maßnahmen („Citizen Voice“) in seinen Berichten auf³³ – beides Aspekte, die in der deutschen Evaluierungspraxis bisher nicht üblich sind.

Für eine starke Relevanz der ICAI in der Entwicklungszusammenarbeit des Vereinigten Königreichs spricht auch ein empirischer Befund: In den vergangenen Jahren referenzierten Referenten des FCDO bereits in 30-40% aller Projekte, die ein Volumen höher als 40 Millionen Pfund umfassten, auch auf Arbeiten und Bewertungen der ICAI, um programmatische Entscheidungen, etwa Aufstockungen oder Neukonzeptionen, zu begründen.³⁴ Ein Indiz dafür, dass es sich hierbei nicht nur um „Gefälligkeitsübungen“ handelt, ist der Umstand, dass auch Programmvorschläge die den Empfehlungen der ICAI *nicht* folgen auf ihre Einordnungen verweisen und sie in der Begründung mitberücksichtigen.³⁵

In Summe zeigt die „UK-Architektur“ das Zusammenspiel von drei komplementären Funktionen unter dem Dach des Chief Economist und des Analysis Directorate: ICAI als unabhängige externe Evaluierung (Rechenschaftspflicht), K4DD als eingebetteten Wissensservice (Evidenzsynthese), und SIEL als Funktionseinheit für strategische Wirkungsevaluierungen (Evidenzgenerierung).

³⁰ ICAI (o. J.). How we work. [\[Link\]](#)

³¹ FCDO (2020). FCDO's review of the Independent Commission for Aid Impact. [\[Link\]](#)

³² ICAI (o. J.). How we work. [\[Link\]](#)

³³ Ibid.

³⁴ Ritchie, E. (2026). How is the UK Aid Watchdog viewed within FCDO? [\[Link\]](#)

³⁵ Ibid.

Relevanz für das BMZ: Die Evidenzarchitektur des Vereinigten Königreichs, inklusive eines Chefökonomens und einer zentralen Analyseeinheit, bietet klare Blaupausen für die Weiterentwicklung eines deutschen Modells der institutionalisierten Arbeit zu Evidenz und Wirksamkeit in der EZ, inklusive klarer Funktionsabgrenzungen. DEval erfüllt eine vielfach zu ICAI vergleichbare Funktion: Unabhängige, extern mandatierte Evaluierung für Rechenschaftslegung. Die Servicestelle hingegen sollte die internen, nachfrageorientierten Funktionen übernehmen – analog zu K4DD (Evidenzaufbereitung) und SIEL (strategische Evaluierungsbeauftragung) – und einen zusätzlichen Fokus auf Kosteneffizienz und -effektivität legen. Die genauere Komplementarität und Funktionsabgrenzung wird in Kapitel 4 dieses Papiers vertieft.

2.3 WEQA – Einheit Qualität und Ressourcen, SECO-WE (Schweiz)

Der Leistungsbereich Wirtschaftliche Zusammenarbeit und Entwicklung innerhalb des schweizerischen Staatssekretariats für Wirtschaft (SECO-WE) verfügt mit dem Ressort³⁶ Qualität und Ressourcen (WEQA) über eine Stelle für unabhängige Evaluationen, Evaluierungsqualität, Meta-Evaluierung und Methodenentwicklung, die Evidenz unmittelbar in das zukünftige Programmdesign des SECO einfließen lässt. Ihre jüngste Weiterentwicklung wurde durch einen Bericht der Geschäftsprüfungskommission des Ständerates (GPK-S) vom November 2023 zur Wirksamkeitsmessung in der internationalen Zusammenarbeit ausgelöst.³⁷ In dessen Folge entwickelte WEQA einen Action Plan 2024–2026, aus dem 2025 der Pilotbetrieb eines neuen Rapid Evidence Review (RER)-Service hervorging: Operative Einheiten innerhalb von SECO-WE können seither gezielte Evidenzsynthesen in Auftrag geben, um strategische Entscheidungen zur Programmplanung vorzubereiten³⁸ (ähnlich dem K4DD-Programm, [siehe 2.2.1](#)).

Neben dem RER-Service wurde in den vergangenen Jahren gemeinsam mit der Direktion für Entwicklung und Zusammenarbeit (DEZA) und einem Dienstleister ein Rapid Evidence Quality Assessment (REQA) entwickelt – ein Methodenraster mit sechs Kriterien (Zweckmäßigkeit der Evaluation, Transparenz, Triangulation, Angemessenheit der Methoden, Stärke der Kausalanalyse, Nützlichkeit der Empfehlungen) – das die methodische Robustheit von Evaluierungen gemäß Standards der Schweizerischen Evaluationsgesellschaft (SEVAL) und der OECD misst.³⁹ Der erste mit REQA erstellte Performance & Quality Report (2023–24) – eine Meta-Analyse der Projektevaluationen – attestierte dem SECO-WE-Portfolio zwar eine insgesamt starke Projektperformance (81 % der evaluierten Projekte mit befriedigender bis sehr guter Bewertung), zugleich aber eine gemischte Evaluierungsqualität: Nur wenige Evaluationen im Zeitraum 2023–24 erreichen sehr hohe Qualität. Über die Hälfte wurde als grundlegend bewertet. 4 von 48 untersuchten Evaluierungen (8%) enthielten eine Attributionsanalyse, 19 (38%) eine Kontributionsanalyse.⁴⁰ Die Befunde verweisen auch explizit auf vergleichbare Ergebnisse

³⁶ In der Schweiz werden mit „Ressort“ nachgeordnete Abteilungen oder Referate bezeichnet.

³⁷ SECO-WE (2025). Performance and Quality Report of the Evaluated Portfolio of SECO's Development Co-operation 2023–2024, S. 2. [\[Link\]](#)

Die Überprüfung der *Wirksamkeit* von Maßnahmen des Bundes durch das Parlament ist in der Schweiz eine in Art. 170 der Bundesverfassung kodifizierte Verfassungsnorm, während eine mögliche deutsche Entsprechung in Art. 114 Abs. 2 GG primär und explizit auf *Wirtschaftlichkeit* zielt.

³⁸ Ibid.

³⁹ Ibid, S.5

⁴⁰ Ibid, S.12

einer Meta-Evaluierung des DEval in Deutschland, wonach z.B. 73% der Evaluierungen dort kaum oder nur teilweise die Angemessenheit und Limitierung der angewandten Evaluierungsmethoden einordnen oder alternative Ansätze diskutieren.⁴¹

Bereits die REQA-Ergebnisse 2025 weisen auf eine deutliche Verbesserung der Evaluierungsqualität in mehreren Bereichen. Während die Stärkung robuster Kausalanalysen weiterhin ein wichtiges Entwicklungsfeld darstellt, ist die Zahl der Kontributionsanalysen (20 von 24 Evaluationen) deutlich gestiegen.⁴²

Mit seinen 2025 aktualisierten Evaluation Guidelines hat die SECO-WE zudem eine über die Qualitätssicherung von Evaluierungen hinausgehende praxisorientierte Empfehlung zur Wahl des passenden kausalen Analyseansatzes entwickelt: Im „Fit for Complexity“-Ansatz wird für jede Wirkungsevaluation zwischen einer Attributions- und einer Kontributionslogik unterschieden. Attribution – ein kontrafaktischer Vergleich etwa über RCTs oder Quasi-Experimente – eignet sich, wenn sich Wirkungen einer Intervention isoliert von Kontextfaktoren messen lassen und ausreichende Fallzahlen für statistische Verfahren vorliegen. Kontribution – mittels theoriebasierter Verfahren wie Contribution Analysis, Realist Evaluation oder Qualitative Comparative Analysis – wird angewendet, wenn mehrere interagierende Faktoren zu einem Ergebnis beitragen oder Wirkungen schwer quantifizierbar sind, etwa bei System- oder Politikwirkungen. Eine Entscheidungstabelle mit Leitfragen, Anforderungen an die Intervention sowie möglichen Ansätzen und Methoden erleichtert die Leistungsbeschreibungen für Evaluierungsaufträge, die in der Folge robustere Erkenntnisse aus den Evaluierungen für Folgeprojekte liefern.⁴³

Relevanz für das BMZ: Der RER-Service ist wie K4DD ein Beispiel für einen nachfrageorientierten Rapid-Evidence-Helpdesk. Darüber hinaus liefert die „Fit for Complexity“-Methodik mit den Kategorien von Attribution und Kontribution einen Ansatz für die konzeptionellen Arbeiten der Servicestelle im BMZ-Portfolio, mit dem die Vielzahl an Projekten und Programmen grundsätzlich nach Wirksamkeitstypen strukturiert werden könnte. Das REQA-Raster bietet zugleich ein Instrument, mit dem die Servicestelle die methodische Qualität des Evaluierungsportfolios des BMZ kontinuierlich prüfen könnte.

2.4 DLL – Development Learning Lab (Norwegen)

Das Development Learning Lab (DLL) ist eine Kooperation des Chr. Michelsen Institute (CMI), der Norwegian School of Economics (NHH/FAIR) und der Universität Bergen (CISMAC). Unter der Leitung von Forschungsdirektor Espen Villanger arbeitet das DLL direkt mit Norad und norwegischen Durchführungsorganisationen zusammen, um evidenzbasierte Erkenntnisse in die Praxis zu überführen.⁴⁴

Ein besonderes Merkmal des DLL ist der enge Austausch zwischen Forschenden und Praktikern in sogenannten *Learning Arenas*. Es handelt sich dabei um strukturierte Formate, in denen Evidenz gemeinsam diskutiert und für die Programmgestaltung aufbereitet wird.

⁴¹ DEval (2022). Meta-Evaluation on the Quality of (Project) Evaluation in German Development Cooperation. S.42 [\[Link\]](#) zitiert in SECO-WE (2025). Performance and Quality Report of the Evaluated Portfolio of SECO's Development Co-operation 2023–2024, S. 13. [\[Link\]](#)

⁴² Im Erscheinen.

⁴³ OECD (2026). Fit for Complexity: Switzerland's Approach to Evaluating Impact. Table 1, S. 2. [\[Link\]](#)

⁴⁴ Development Learning Lab (o. J.). More Learning, Better Outcomes. [\[Link\]](#)

Für 2025 wurden thematische Schwerpunkte zur wirtschaftlichen Stärkung von Frauen (*Women's Economic Empowerment*), Klimaanpassung (*Climate Adaptation*) und Arbeitsmärkten (*Jobs Network*) gesetzt.⁴⁵ Das DLL arbeitet dabei entlang des gesamten Projektzyklus: Wissenslücken und Bedarfe werden im kontinuierlichen Dialog mit Partnern identifiziert, und die Forschenden etablieren schnelle Feedback-Schleifen, damit wichtige Erkenntnisse zeitnah in die Praxis einfließen.⁴⁶

Ferner wurden drei verschiedene Formate an Wissensübersichten (Knowledge Reviews) etabliert:

- **Knowledge Briefs:** Kompakte Dokumente (3-4 Seiten) mit unmittelbar handlungsweisenden Schlussfolgerungen,
- **Evidence Reviews:** Ausführlichere Evidenzübersichten für Programm- und Projektkonzeptionen,
- **Systematic Reviews:** Umfassende Analysen angelehnt an rein akademische Arbeiten, die die anderen Formate fundieren und vor allem bei übergeordneten strategischen Entscheidungen unterstützen sollen.

2024 veröffentlichte DLL acht Evidence Reviews⁴⁷, 2025 drei weitere⁴⁸. Zusätzlich werden gegenwärtig neun Forschungsprojekte mit NGOs und Regierungen durchgeführt. Diese wenden überwiegend RCTs an, oder kombinierte rigorose Evidenzdesigns. Ein Forschungsvorhaben zu Interventionen gegen weibliche Genitalverstümmelung und Kinderehe umfasst beispielsweise Effekte auf 3.100 Eltern und 12.000 Kinder.⁴⁹

Auch hier geht die fachliche Arbeit zur Konsolidierung von Evidenz mit Bestrebungen zum Kapazitätsaufbau im Ökosystem einher: In Zusammenarbeit mit DLL und der International Initiative for Impact Evaluation (3ie) wurde der erste „Impact Evaluation Incubator“ ins Leben gerufen und verstetigt, ein dreitägiges intensives Fortbildungsformat, in dem Programmverantwortliche lernen, Wirkungsmessungen in ihre Vorhaben zu integrieren.⁵⁰

Eine aktuelle Evaluierung zu Norads Wissensnutzung im Portfoliomanagement zeigt auch die Herausforderungen: In kurzer Zeit wurden substanzielle Fortschritte erzielt, es verbleiben aber auch noch Baustellen: Wissen fließt vor allem in Design und Problemanalyse ein, kaum dagegen in laufende Projektsteuerung und Anpassung; Monitoring- und Evaluierungsdaten werden seltener genutzt als Forschungs- und Erfahrungswissen; und es mangelt weiter an portfolioübergreifenden Kompetenzen zu Monitoring und Evaluierung.⁵¹

Politische Rückendeckung erhielten die Aktivitäten des DLL auch durch die explizite Formulierung im norwegischen Staatshaushalt 2023, dass die norwegische Entwicklungszusammenarbeit eine hohe Qualität aufrechterhalten und wissensbasiert sein soll.⁵² Erst kürzlich erhielt DLL zudem weitere Mittel von Norad, um seine Aktivitäten in den

⁴⁵ DLL (2025). Annual Report. [\[Link\]](#)

⁴⁶ Ibid.

⁴⁷ DLL (2024). Annual Report. [\[Link\]](#)

⁴⁸ DLL (2025). Annual Report. [\[Link\]](#)

⁴⁹ Ibid.

⁵⁰ DLL (2024). Annual Report. S. 10 [\[Link\]](#)

⁵¹ Lloyd, R., Tjønneland, E. et al. (2024). Evaluation of Norad's use of knowledge in portfolio management. [\[Link\]](#)

⁵² Norwegian Task Force on evidence-based aid (2023). Impact evaluations in Norwegian development cooperation – Opportunities, challenges and proposed actions. S. 3 [\[Link\]](#)

nächsten 5 Jahren zu einem Knowledge Centre for Development Cooperation (Wissenszentrum für Entwicklungszusammenarbeit) auszubauen.⁵³

Relevanz für das BMZ: Das DLL-Modell ist besonders relevant für die Zusammenarbeit zwischen BMZ und seinen Durchführungsorganisationen GIZ und KfW. Die Learning Arenas bieten ein konkretes Format, um z.B. die im Reformplan geforderten Standardwirkungsmodelle partizipativ zu entwickeln und in die Programm- und Projektplanung zu übernehmen. Auch das BMZ könnte hier zu Beginn strategisch besonders relevante Themenschwerpunkte setzen. Der Impact Evaluation Incubator ist ein sofort adaptierbares Format für den Kapazitätsaufbau in BMZ, GIZ und KfW und könnte zum Beispiel die Translation aus Standardwirkungsmodellen in Projektplanung unterstützen.

2.5 ACRES – Center for Rapid Evidence Synthesis (Uganda)

Das Center for Rapid Evidence Synthesis (ACRES) unter der Leitung von Dr. Rhona Mijumbi-Deve ist eine in Kampala ansässige Nichtregierungsorganisation zur Wissens- und Evidenzvermittlung für Entscheidungsträger im lokalen Mehrebenensystem.⁵⁴ Die langjährige Arbeit zu angewandten Evidenzsynthesen erstreckt sich über mehrere Sektoren: Öffentliche Gesundheit, Bildung, Makroökonomie, Klimawandel und weitere.

Besonders bemerkenswert ist der Einsatz KI-gestützter Tools: ACRES hat in den vergangenen drei Jahren die Digital Rapid Evidence Synthesis (Dr-RES) Plattform⁵⁵ entwickelt, die über 4000 Evidenzsynthesen aus 16 Jahren schneller, übersichtlicher und verständlicher macht und wertvolle Blaupausen für eine moderne Evidenzinfrastruktur des BMZ und seiner Durchführungsorganisationen liefern kann. Die eigenständige Entwicklung maßgeschneiderter digitaler EIDM-Tools (Evidence-Informed Decision-Making) in Ländern des Globalen Südens löst diese aus der Abhängigkeit von Bewertungsnormen und Standards aus dem Globalen Norden und ermöglicht neue Kooperationspotentiale partnerschaftlicher Nord-Süd-Wirkungsevaluierung.

ACRES koordiniert zudem das Evidence to Policy (E2P) Training Programme im Rahmen der Africa LEEPS-Initiative, das in fünf afrikanischen Ländern eine neue Generation von EIDM-Führungskräften ausbildet.⁵⁶ Der Ansatz betont, dass technische Kompetenzen allein nicht ausreichen: Sozialkompetenzen, etwa das Navigieren von Machtdynamiken oder Kollaborationsfähigkeiten zwischen Forschungs- und Politiklogiken sind ebenso zentral.

Relevanz für das BMZ: ACRES ist besonders relevant als Beispiel für Evidenzinfrastruktur, die direkt im Globalen Süden verankert ist. Die Servicestelle könnte perspektivisch Partnerschaften mit ACRES und vergleichbaren Institutionen aufbauen, um lokale Evidenz besser in die Programmgestaltung einzubeziehen – im Einklang mit dem Reformplan-Ziel der stärkeren Partnerorientierung. Der KI-gestützte Ansatz von ACRES bietet zudem Anknüpfungspunkte für die Entwicklung von KI-gestützten Verfahren, wie sie auch das BMZ entwickeln möchte, und die die Servicestelle perspektivisch in ihre Fachverfahren integrieren könnte.

⁵³ CMI (2026). CMI awarded funding for new Knowledge Centre to strengthen evidence in aid. [\[Link\]](#)

⁵⁴ ACRES (2026). Transforming decision-making in Africa. [\[Link\]](#)

⁵⁵ ACRES (2026b). Digital Evidence Hub Dr-RES. [\[Link\]](#)

⁵⁶ ACRES (o. J.). The Evidence to Policy (E2P) Training Programme. [\[Link\]](#)

2.6 PACE – Pan-African Collective for Evidence (Südafrika)

Das Pan-African Collective for Evidence (PACE) ist eine in Johannesburg ansässige Nichtregierungsorganisation unter der Leitung von Laurenz Mahlanza-Langer, die sich auf die Unterstützung evidenzbasierter Politikgestaltung in Afrika spezialisiert hat.⁵⁷ Mit dem Africa Evidence Network (AEN) unterhält PACE das laut eigener Aussage größte afrikanische Netzwerk für evidenzbasierte Entscheidungsfindung. Unterschiedliche Arbeitsstränge zu Evidenzgenerierung und -synthese unterstützen Regierungen bei der Integration von Evidenz in Politikprozesse.

PACEs systematische Übersichtsarbeit (Systematic Review) *The art and science of promoting evidence-informed decision-making: A global living evidence map*⁵⁸ ordnet verschiedene Interventionen nach ihrem Potential zur Förderung evidenzbasierter Entscheidungsfindung. Die Arbeit identifiziert sechs Wirkmechanismen zu einer verbesserten Evidenznutzung – von der Stärkung individueller Fähigkeiten bis zur Veränderung institutioneller Anreize. Aktuell führt PACE im Auftrag des FCDO eine weitere systematische Übersichtsarbeit zu der Frage durch, welche Maßnahmen die Evidenznutzung bei Entscheidungsträgern am wirksamsten verstärken. Das Hinzuziehen dieser Ergebnisse würde einer Servicestelle Evidenz zu innerer Kohärenz verhelfen: Sie richtet den Evidenzanspruch damit auch auf sich selbst und ihre eigene Ausgestaltung und Arbeitsweise.

Schließlich koordiniert PACE das *Responsive Evidence System for African Policy Needs* (REAP)-Projekt, das eine skalierbare Online-Plattform für den Kapazitätsaufbau von Evidenzeinheiten in Afrika entwickelt und damit Evidenzeinheiten in fünf afrikanischen Ländern unterstützt.⁵⁹

Relevanz für das BMZ: PACE bietet Expertise in der Frage, wie Evidenznutzung in Organisationen und Regierungen auf dem afrikanischen Kontinent institutionalisiert werden kann. Diese Meta-Perspektive – die Wissenschaft der Evidenznutzung selbst – ist für den Aufbau der Servicestelle hochrelevant: Sie kann helfen, die richtigen institutionellen Anreize und Formate zu wählen, um Evidenz tatsächlich in Entscheidungen einfließen zu lassen. Zudem könnten die REAP-Erfahrungen auch ein Modell für den Kapazitätsaufbau von Evidenzeinheiten in BMZ-Partnerländern sein.

2.7 DIME und SIEF – Development Impact Evaluation und Strategic Impact Evaluation Fund (Weltbank)

DIME, die Einheit für Wirkungsevaluierungen in der Weltbank, hat über die letzten 20 Jahre eine Vielzahl an Datenprodukten und Wirkungsevaluierungen zur effektiveren Ausgestaltung von Politikmaßnahmen in Partnerländern entwickelt. Offen wird hier eine Strategie des „Ausprobierens und Anpassens“ (trial-and-adopt) angewandt, bei der auf Arbeitsebene in

⁵⁷ PACE (2026). Pan-African Collective for Evidence. [\[Link\]](#)

⁵⁸ WHO (2024). The art and science of promoting evidence-informed decision-making: a global living evidence map. [\[Link\]](#)

⁵⁹ REAP (2026). Responsive Evidence System for African Policy Needs. [\[Link\]](#)

einzelnen Projekten Wirkungshypothesen aufgesetzt und laufend rigoros getestet werden. Daraus abgeleitete Learnings bezüglich Kosteneffektivität und wirksamen Projektdesigns werden anschließend im Portfolio der Weltbank skaliert.⁶⁰

Viele der Mikrodaten, die für die Wirkungsevaluierungen verwendet wurden, sind öffentlich zur Replizierung oder weiteren Forschung in einem Datenkatalog bereitgestellt.⁶¹ Übergeordnetes Prozess- und Praxiswissen wird außerdem öffentlich in einer DIME Wiki geteilt.⁶² Zusätzlich wurden in den vergangenen Jahren KI-gestützte Lösungen für die operative Ebene entwickelt. Das Projekt ImpactAI⁶³ bereitet jeweils aktuelle und rigorose Evidenz zu EZ-Maßnahmen auf – inklusive Visualisierung von Daten – und ist auf Anfrage auch für Fachnutzer außerhalb der Weltbank verfügbar. Zudem entwickelt DIME themenspezifische KI-Lösungen, etwa Dashboards zur Echtzeiterfassung von drohenden Hungerkrisen oder Sprachmodelle zur Erfassung von Hate Speech.⁶⁴

SIEF ist ein von der Weltbank koordinierter und geleiteter Multigeber-Fonds zur Finanzierung von Wirkungsevaluierungen in den Sektoren Gesundheit, Bildung, Soziale Sicherung und Beschäftigung, mit Fokus auf randomisierten Kontrollstudien (RCT) und quasi-experimentellen Studien. Seit 2011 hat SIEF mehr als 100 solcher rigoroser Wirkungsevaluierungen in mehr als 40 Ländern mit niedrigem und mittlerem Einkommen finanziert und darüber hinaus eine Vielzahl an Workshops und Wissensmaterialien zur Kapazitätsstärkung im Bereich Evidenz und Wirksamkeit entwickelt.⁶⁵

Aktuell wird die frei zugängliche Online-Datenbank IDEAL (Impact Data and Evidence Aggregation Library) eingerichtet, in der standardisierte Wirkungsabschätzungen aus randomisierten kontrollierten Studien gebündelt werden.⁶⁶ Nutzer können darin relative Effektgrößen für eine Vielzahl von Ergebnissen und unterschiedlicher politischer Maßnahmen über Länder und Kontexte hinweg suchen, vergleichen und visualisieren.

Übergeordnet bietet die Weltbank mit ihrer Scorecard⁶⁷ eine aggregierte, öffentliche Übersicht ihrer Wirkungen und Standardindikatoren an und stellt diese auch per Schnittstelle zur Verfügung. Die hier berichteten Werte sind unter anderem auf den Evidenzierungsarbeiten von DIME und SIEF abgestützt. Die Indikatoren verbinden hier retrospektive und prospektive Werte und kontextualisieren sie gegenüber zentralen, weltweiten Entwicklungs Herausforderungen.

Relevanz für das BMZ: Die Servicestelle könnte perspektivisch in ähnlicher Weise wie DIME Datenanalyse und Wirkungsorientierung bündeln, und das ressortinterne Wissen zu Evidenz und Wirksamkeit in niedrigschwelliger Form konsolidieren und zur Verfügung stellen. Sie könnte eine ähnliche Datenbank wie IDEAL aufbauen, die sowohl als Transparenzinstrument und Service für Referenten, als auch als öffentliches Gut für die Fachcommunity und die breite Öffentlichkeit dient. Stetiges Ziel muss es sein, Wirksamkeitsdaten standardisiert und vergleichbar zu machen, auch als Grundvoraussetzung für die im BMZ-Reformplan geforderte wirkungsorientierte Mittelallokation.

⁶⁰ Weltbank (o.J.). Development Impact Group - Driving Policy Impact. [\[Link\]](#)

⁶¹ Weltbank (o.J.). Microdata Library. Development Impact Evaluation (DIME). [\[Link\]](#)

⁶² Weltbank (o.J.). Dimewiki. [\[Link\]](#)

⁶³ Weltbank (o.J.). ImpactAI - Better Evidence. Better Decisions. [\[Link\]](#)

⁶⁴ Weltbank (o.J.). Development Impact Group – Artificial Intelligence. [\[Link\]](#)

⁶⁵ Weltbank (o. J.). The Strategic Impact Evaluation Fund (SIEF). [\[Link\]](#); Microdata Library. [\[Link\]](#)

⁶⁶ Weltbank (o. J.). Brief - Impact Data and Evidence Aggregation Library (IDEAL). [\[Link\]](#)

⁶⁷ Weltbank (o.J.). The Worldbank Group Scorecard. [\[Link\]](#)

2.8 AIL - Airbel Impact Lab des International Rescue Committee (IRC)

Das International Rescue Committee hat in den vergangenen Jahren das Airbel Impact Lab als internes und eigenständiges Forschungs- und Innovationsteam aufgesetzt und damit seine Arbeiten zu Evidenzgenerierung, Wirksamkeitsmessung und Kosteneffektivitätsschätzung im humanitären Sektor stark ausgeweitet. Laut eigener Aussage ist das IRC damit für 19% aller Wirkungsevaluierungen im humanitären Sektor verantwortlich, obwohl es nur 3% seines Budgets umsetzt.⁶⁸

Das „Best Use of Resources“ (BUR) Team fokussiert als Untereinheit Kosteneffizienz- und Kosteneffektivitätsanalysen für die Projekte des IRC: In den letzten 10 Jahren wurden über 400 Kosteneffizienz- und 24 Kosteneffektivitätsanalysen in Verbindung mit Wirkungsevaluationen erstellt. Der wachsende Korpus steht in einer öffentlichen Online-Bibliothek zur Verfügung.⁶⁹

Eine interaktive Übersicht der Wirkungshypothesen und Evidenzen (interactive Outcomes and Evidence Framework)⁷⁰ mit denen das IRC arbeitet, sowie das webbasierte Tool „Dioptra“, mit dem Programmplaner eine umfassende Kostenanalyse vollziehen können,⁷¹ sind komplementäre Bestrebungen, die Evidenz- und Wirkungsanalysen unmittelbar in Arbeitsschritte des operativen Bereichs überführen. Im Auftrag von FCDO setzt das Airbel Impact Lab gegenwärtig auch das Programm Education Research in Crisis and Conflict (ERIC) um, das Evidenzlücken und kontextsensible, herausragend effektive Ansätze für den Zugang zu Bildung in Krisenregionen identifiziert.⁷²

Relevanz für das BMZ: Die Servicestelle könnte insbesondere die Kosteneffizienz- und Kosteneffektivitätsanalysen als praxiserprobte Referenzpunkte für ihre eigenen entsprechenden Arbeiten nehmen. Durch die Zusammenarbeit mit dem IRC in der Vergangenheit im Rahmen der Übergangshilfe des BMZ gibt es auch etablierte Schnittstellen und Erfahrungswerte, die für einen zusätzlichen Wissenstransfer genutzt werden können.

⁶⁸ IRC (o. J.). Research and Innovation at the International Rescue Committee. [\[Link\]](#)

⁶⁹ IRC(o. J. b). Explore Cost Analysis at the IRC. [\[Link\]](#)

⁷⁰ IRC (2023). Interactive Outcomes and Evidence Framework. [\[Link\]](#)

⁷¹ Dioptra (o. J.). Dioptra - Rapid, rigorous cost-efficiency analysis. [\[Link\]](#)

⁷² IRC (o. J. c). Education Research in Conflict and Protracted Crisis (ERICC). [\[Link\]](#)

2.9 Vergleichende Übersicht

Institution/ Projekt	Land/ Ort	Kernfunktion	Kennwerte	Relevante Aspekte für BMZ
USAID OCE	USA	Kosten-Effektivitäts-Analysen, Haushaltssteuerung	30+ Mitarbeitende; 1,7 Mrd. USD umgeschichtet	Umfassendster Referenzpunkt mit größter Skala, formelles Mandat, operative Einbettung
ECDO	UK	Bündelung von Analytik und Evidenz-Checks	15 Jahre Erfahrung mit Kosten-Nutzen-Analysen, interne Best-Buys-Listen	Langer Erfahrungszeitraum, operative Einbettung, Aufteilung der Funktionen
K4DD	UK	Rapid-Evidence-Re views, Helpdesk, Lernformate	200+ Reviews/Jahr; 6 Konsortialpartner	Helpdesk-Modell, Konsortialansatz
WEQA	CH	Rapid-Evidence-Ser vice; Rahmenwerk für Evaluations- und Evidenzqualität	Pilot seit 2025 (erste 3 RER); REQA mit 6 Kriterien	Helpdesk-Modell, übertragbarer Methoden- rahmen zur Wahl der Evaluierungslogik
SIEL	UK	Rigorese Wirkungsevaluierun gen (RCTs), Kapazitätsaufbau	6-Jahres-Programm, Kooperation zwischen Innovations for Poverty Action (IPA) & J-PAL	Strategische Evaluierungsbeauftra gung, Learning Agendas
ICAI	UK	Unabhängige Evaluierung, parlamentarische Kontrolle	3 Commissioners, 10 weitere Mitarbeitende	Komplementarität zu DEval, an Parlament gerichtete Wirkungskommunikation, Prüfung Kosteneffizienz
DLL	NOR	Praxisnahe Forschung, Learning Arenas, Incubator	3 Learning Networks, 8 Evidence Reviews in 2024, 9 rigorose Wirkungsevaluierungen 2025	GIZ/KfW-Kooperation, Praxistransfer, horizontale Evidenzgenerierung und -nutzung
ACRES	UG	Rapid-Evidence-Sy nthesis, KI-Tools, Evidence-to-Policy- Training	Über 4000 lokal relevante Evidenzsynthesen aus 16 Jahren digital aufbereitet	Süd-Partnerschaften, KI-Nutzung
PACE	ZAF	Evidenz zur Evidenznutzung, Netzwerkkoordinatio n	Koordinierung und Unterstützung von 5 Evidenz- einheiten in Afrika	Süd-Partnerschaften, Meta-Expertise zur Evidenznutzung
DIME/SIEF	Welt- bank	Wirkungsevaluierun g, Open-Data-Plattfor m	100+ rigorose Wirkungsevaluierungen in 40+ Ländern, umfassende Scorecard	Strategische Wirkungsevaluierungen (inkl. RCTs), Transparenztool, Datenbank-Vorbild
AIL	IRC	Wirkungsevaluation en und detaillierte Kostenanalysen	400 Kosteneffizienz- und 24 Kosteneffektivitätsanalys en in 10 Jahren	Praxiserprobte Kosteneffizienz- & Kosteneffektivitätsanalysen; Beispiele für web-basierte Einbettung der Ergebnisse im operativen Bereich

3. Empfohlene Mandate und Funktionen der Servicestelle

Auf Grundlage der internationalen Erfahrungen empfehlen wir der Servicestelle **neun Funktionen**:

Nr.	Funktion	Outputs und intl. Beispiele
1	Wirkungsorientierte Haushaltssteuerung Zuarbeit an die Leitungsebene: vergleichende Portfolioanalyse, Identifikation von „Smart Buys“ und rechtzeitiger Input für den Haushaltsprozess.	Analysen und Voten zum Haushaltsprozess; Smart-Buys-Übersichten USAID OCE
2	Verbindlicher Evidenz- und Kosten-Effektivitätscheck Standardschritt in der Programmkonzeption, proportional zur Projektgröße; Blick auf Evidenzlage, Alternativen und Referenzwerte.	Standardisierte Checklisten für Programmplanung; Referenzwerte FCDO
3	Konzeptionelle Arbeiten und Portfoliobericht Identifikation zusätzlicher Wirkungspotentiale im BMZ-Portfolio; Testung sektorimmanenter und -übergreifender Vergleichsmetriken.	Portfolio-Wirksamkeitsbericht; Vergleichsmetriken USAID OCE , FCDO , WEQA
4	Rapid-Evidence-Helpdesk Schnelle Evidenzsynthesen für Referate auf Abruf, unterstützt von digitalen Lösungen.	Rapid Evidence Reviews; Knowledge Briefs K4DD , WEQA
5	Kapazitätsaufbau und Lernkultur Ressortinterner Aufbau von Evidenzkompetenz; offene Fehlerkultur und praxisnahe Lernformate.	Schulungen; Impact Evaluation Incubator; Learning Agendas DLL (Norad), BRAC
6	Strategische Beauftragung von Wirkungsevaluierungen Identifikation von Evidenzbedarfen und gezielte Vergabe, koordiniert mit dem Evaluierungsreferat des BMZ.	Learning Agendas; beauftragte Wirkungsevaluierungen SIEL
7	Vernetzung und internationaler Austausch Kontinuierlicher Austausch mit internationalen Evidenzeinheiten und Partnern zu Standards und Initiativen.	Mitwirkung in Netzwerken; gemeinsame Evidenzinitiativen Smart Buys Alliance
8	Aufbau einer Wirksamkeitsdatenbank Mitwirkung an global koordinierter Evidenz; perspektivisch eine öffentliche Datenbank, gestützt auf Forschungspartner.	Wirksamkeitsdatenbank; Beiträge zu offenen Datenplattformen Weltbank (IDEAL-Plattform)
9	Strategische Kommunikation von Wirksamkeit Gemeinsam mit dem Kommunikationsreferat die Wirksamkeit der deutschen EZ sichtbar machen und Vertrauen stärken.	(Standard-)Indikatoren; Beiträge zur Öffentlichkeitsarbeit Weltbank Scorecard

Diese Funktionen sind so geordnet und konzeptionalisiert, dass sie primär auf die Steigerung der Wirksamkeit des BMZ-Portfolios ausgerichtet sind und die bestehenden Strukturen der deutschen EZ stärken, nicht duplizieren. Entsprechend verweist ein kurzer Exkurs am Ende dieses Kapitels auch auf bereits bestehende Praktiken einer evidenzbasierten Steuerung in der deutschen EZ hin, auf denen aufgebaut werden kann.

3.1 Funktion 1: Wirkungsorientierte Haushaltssteuerung

Die Servicestelle sollte direkt der Leitungsebene für eine Evidenz- und Wirkungsorientierung in der Haushaltssteuerung und Mittelpriorisierung zuarbeiten. Konkret bedeutet dies:

- **Vergleichende wirkungsorientierte Portfolioanalyse:** Systematische Analyse des BMZ-Portfolios, um bewährte Ansätze mit dem besten Verhältnis von Wirksamkeit, Wirtschaftlichkeit und Umsetzungsfähigkeit zu identifizieren sowie weniger wirksame Maßnahmen zu benennen, die gemäß Reformplan „schrittweise zurückzuführen“ sind⁷³.
- **Identifikation von „Smart Buys“:** Systematische Aufbereitung der internationalen Evidenz zu herausragend wirksamen und kosteneffektiven Ansätzen („bewährte Ansätze“ im Sinne des Reformplans), differenziert nach Sektoren und Zielen des BMZ. Hierfür kann die Servicestelle zum Beispiel auf die Arbeit der Smart Buys Alliance⁷⁴ zurückgreifen.
- **Inputfunktion für Haushaltsprozesse:** Rechtzeitige Bereitstellung von Wirksamkeitsanalysen für den Haushaltsprozess.

Die einschlägigen Vorbilder für diese klar ressortimmanente Funktion sind das USAID OCE ([siehe 2.1](#)) und FCDO ([siehe 2.2](#)). Für das BMZ könnten selbst moderate Umschichtungen von 5–10 Prozent des Einzelplan 23 auf der Grundlage von Wirksamkeitskriterien, sowohl innerhalb von Titeln als auch im Gesamtportfolio, einen erheblichen Zugewinn an Wirkungen der deutschen EZ bedeuten.

3.2 Funktion 2: Verbindlicher Evidenz- und Kosten-Effektivitäts-Check

Die Servicestelle sollte einen verbindlichen Evidenz- und Kosten-Effektivitäts-Check als Standardschritt in der Programmkonzeption und Beauftragung etablieren. Dieser Check sollte proportional zur Projektgröße gestaltet werden:

- Für große Vorhaben (beispielsweise ab 20 Mio. Euro) eine vertiefte Analyse
- Für kleinere Vorhaben eine standardisierte Checkliste.

Bestehende Prozesse, wie die interne Planung und Allokation (iPA) oder die maßnahmenscharfe Planung und Prüfmissionen, sollten entsprechend angepasst und vereinheitlicht werden. Grundsätzlich sollte bei der Konzeption neuer Programme systematisch geprüft werden:

- Welche Evidenz liegt für den geplanten Interventionstyp vor?
- Wie verhält sich die erwartete Wirkung im Vergleich zu alternativen Ansätzen im gleichen Sektor?

⁷³ BMZ (2026). Zukunft zusammen global gestalten. S.24. [\[Link\]](#)

⁷⁴ Smart Buys Alliance (2026). The Smart Buys Alliance. [\[Link\]](#)

- Welche Smart Buys können als Referenzwerte dienen?
- Welche Referenzgruppen und Kennwerte sollten bei neuen Programmen als Vergleichsgrundlage definiert werden?

Die Nutzung und Deliberation bezüglich Evidenz sollte sowohl in formellen Prozessen verankert als auch auf sozialer Ebene kultiviert werden (siehe Funktion 5).

Zu Evidenz- und Kosteneffektivitätsanalysen gibt es neben den Praxisbeispielen des IRC ([siehe 2.8](#)) eine wachsende wissenschaftliche Literatur⁷⁵ und Vorarbeiten anderer Geberländer, die zur Ausgestaltung in der Servicestelle Evidenz genutzt werden können ([siehe OCE, 2.1](#); oder [FCDO, 2.2](#)).

3.3 Funktion 3: Konzeptionelle Arbeiten und Portfoliobericht zu Evidenz- und Wirksamkeit im BMZ

Die Servicestelle sollte kontinuierlich die ressortinterne konzeptionelle Arbeit zu Evidenz und Wirksamkeit vorantreiben und dem BMZ in evidenzbasierter Weise darlegen, welche Wirkungspotentiale in seinem Portfolio bestehen.

Das Beispiel des OCE ([siehe 2.1](#)) macht vor, wie man hierzu erste gezielte Arbeiten bereits mit einem kleinen Team und in nachfrageorientierter und priorisierter Weise beginnen kann: Identifiziert werden müssen die Programme mit den größten Mittelvolumina; in deren fachlicher Domäne Evidenzlücken bestehen, oder deren aktuelle Konzeption ein relevantes Delta zum wissenschaftlichen Stand der Evidenz aufweisen; und deren Referate und fachliche Partner eine Eigenmotivation zeigen, die Arbeit zu Evidenz und Wirksamkeit gemeinsam zu intensivieren.

Daraus sollte sich regelmäßig ein aggregierter Portfoliobericht zu Evidenz und Wirksamkeit im BMZ speisen. Dieser vergleicht die Wirksamkeit verschiedener Ansätze *innerhalb* eines Sektors im laufenden Portfolio. Auch die erzielten Wirkungen *zwischen* Sektoren sollten fundiert diskutiert werden, um zum Beispiel die in mehreren Reformprozessen angestrebte geographische und thematische Fokussierung der bilateralen EZ zu unterstützen⁷⁶. Im Detail sollte der Portfoliobericht zu Evidenz und Wirksamkeit:

- Die Wirksamkeit der wichtigsten Programmlinien anhand standardisierter Indikatoren vergleichen.
- Explizit benennen, welche Ansätze ein überzeugendes Verhältnis von Wirksamkeit, Wirtschaftlichkeit und Umsetzungsfähigkeit aufweisen.
- Konkrete Empfehlungen für die Stärkung bewährter Ansätze und die Rückführung und Ausphasierung weniger wirksamer Maßnahmen ableiten.
- Als Grundlage für den Haushaltsprozess und die Kommunikation gegenüber dem Parlament dienen.

Empfohlen wird ein Ampelsystem, wie es zum Beispiel von ICAI ([siehe 2.2.3](#)) praktiziert wird. Auch Kosteneffizienz- und Kosteneffektivitätsaufstellungen sollten Bestandteil des

⁷⁵ Vrgl. statt vieler Glandon, D., Fishman, S., Tulloch, C., et al. (2023). The State of Cost-Effectiveness Guidance: Ten Best Resources for CEA in Impact Evaluations. *Journal of Development Effectiveness*, 15(1), 5–16. [\[Link\]](#)

⁷⁶ Vrgl. Wencker, T. (2022). Die Verteilung von Mitteln für die deutsche öffentliche Entwicklungszusammenarbeit. DEval. [\[Link\]](#)

Berichts sein und könnten mit den ex-ante Kosteneffektivitätschecks ([siehe 3.2](#)) abgeglichen werden. Hierfür ist auch ein Blick auf die Beispiele des IRC ([siehe 2.8](#)) naheliegend.

Neben den bestehenden Standardindikatoren, sollte auch die im BMZ-Reformplan genannte Einführung von Standardwirkungsmodellen durch GIZ und KfW in den Wirksamkeitsbericht integriert werden. Ferner sollte er eine strukturierte Darlegung der unterschiedlichen Evidenzlogiken im BMZ-Portfolio beinhalten, ähnlich z.B. der Arbeiten der SECO in der Schweiz ([siehe 2.3](#)). Systematische Wirkungsevaluierungen sollten sowohl auf Programm- und Modulebene, als auch auf Länderebene ausgebaut und einbezogen werden.⁷⁷

Mittelfristig kann über diese Funktion in Zusammenarbeit mit dem Datenlabor des BMZ auch eine Stärkung der ressortinternen Analytics-Fähigkeiten stattfinden, also die Auswertung von nicht nur statischen und herkömmlichen, sondern auch unkonventionellen und dynamischen Datenquellen⁷⁸ und Echtzeitdaten. KI-gestützte Verfahren und neue Formen der Datenauswertung können zum Beispiel präzisere, kosteneffizientere und effektivere Präventionsmaßnahmen⁷⁹ in Krisen- und Entwicklungskontexten ermöglichen und fördern damit die Handlungsfähigkeit des BMZ.

Neben der dadurch noch stärker berücksichtigten zeitlichen Dimension von Maßnahmen, ist diese Funktion der Servicestelle Evidenz ein struktureller Beitrag zur sektorübergreifenden Priorisierung. Diese findet angesichts sinkender ODA-Mittel bereits jetzt ohnehin und notgedrungen statt. Der Portfolio-Wirksamkeitsbericht sollte sie um die Anwendung von sektorübergreifenden Vergleichsmetriken ergänzen (z.B. Kosteneffektivität, QALYs⁸⁰ oder WELLBYs⁸¹), zu denen es jeweils eine wachsende Forschungsliteratur gibt. Grenzen der Messbarkeit sollten durch den Bericht sowohl getestet, als auch ausgewiesen werden, und damit fachliche und politische Aushandlungsprozesse konstruktiv begleitet werden.

3.4 Funktion 4: Rapid-Evidence-Helpdesk

Nach dem Vorbild von K4DD ([siehe 2.2.1](#)) und WEQA ([siehe 2.3](#)) sollte die Servicestelle einen Rapid-Evidence-Service für BMZ-Referenten anbieten. Über ein Funktionspostfach und perspektivisch eine digitale Plattform können Mitarbeitende schnelle Evidenzsynthesen zu spezifischen Programmfragen anfordern. Für das BMZ könnte dieses Format über bestehende und neue Partnerschaften mit IDOS, GIGA, Universitäten und Instituten mit Evidenzexpertise realisiert werden.

Das Rapid-Evidence-Helpdesk ist zudem eine passende Schnittstelle zum Datenlabor des BMZ, denn jede Evidenz und Wirksamkeitsanalyse ist auf gute Datengrundlagen angewiesen. Umgekehrt sind Erfahrungswerte und Anfragen aus dem Rapid-Evidence-Helpdesk gute Bedarfsorientierungen für das Datenlabor. Ferner ist das Rapid-Evidence-Helpdesk auch ein hervorragender Anwendungsbereich für BMZ-interne Bestrebungen zu digital- und KI-gestütztem Arbeiten, wie sie beispielhaft durch ImpactAI der

⁷⁷ Vrgl. BRH (2023). Bericht über die Evaluierung von Maßnahmen der Entwicklungszusammenarbeit. [\[Link\]](#)

⁷⁸ Zable, A., Verhulst, S. (2025). Monitoring the Re-Use and Impact of Non-Traditional Data. [\[Link\]](#)

⁷⁹ Siehe statt vieler Jing, W., Tasci, K., Ergun, Ö., Vosti, S., Webb, P. (2025) Enhancing Cost-Efficiency and Effectiveness in USAID's Food Aid Supply Chain Operations in Ethiopia. [\[Link\]](#)

⁸⁰ Feng, X., Kim, D. D., Cohen, J. T., et al. (2020). Using QALYs versus DALYs to measure cost-effectiveness: how much does it matter? [\[Link\]](#)

⁸¹ Frijters, P., Krekel, C., Sanchis, R., Santini, Z. I. (2024). The WELLBY: a new measure of social value and progress. [\[Link\]](#)

Weltbank ([siehe 2.7](#)) oder der Digital Rapid Evidence Synthesis (Dr-RES) Plattform von ACRES ([siehe 2.5](#)) aufgezeigt werden.

3.5 Funktion 5: Kapazitätsaufbau und Lernkultur

Die Servicestelle sollte systematisch den Kapazitätsaufbau für evidenzbasierte Entscheidungsfindung in der gesamten Organisation fördern. Das norwegische Modell des Impact Evaluation Incubator ([siehe 2.4](#)) und das SIEL-Modell ([siehe 2.2.2](#)) mit seinen maßgeschneiderten Schulungen bieten konkrete Beispielformate. Darüber hinaus sollte die Servicestelle einen leitenden Beitrag zum Aufbau einer ressortinternen Fehler- und Lernkultur leisten, in der die Identifikation weniger wirksamer Ansätze nicht als Scheitern, sondern als Grundlage für bessere zukünftige Entscheidungen verstanden wird. Neben einer verbesserten *technischen* Ausstattung des Wissensmanagements sind hierfür vor allem auch *soziale* Lernformate vonnöten – wie auch BMZ-Mitarbeitende selbst betonen.⁸²

In Norad ist hierfür beispielsweise das Format eines „Speaker’s Corner“ institutionalisiert, bei dem offen im Büro ein konkretes Problem gestellt wird und mögliche Lösungen diskutiert werden. Breit angelegte Deliberationen über die Organisationsziele, auch ohne förmliche Rahmen, sowie Gespräche über Misserfolge und Fehler, sind ebenfalls etabliert.⁸³

Ein weiteres Beispiel hierzu ist die nichtstaatliche Entwicklungsorganisation BRAC, die den Anspruch „aus Fehlern zu lernen“ sehr wörtlich nimmt. Sie erstellt und veröffentlicht bereits seit mehreren Jahren jährlich einen eigenen „Fehlerbericht“ (Failure Report), um ihr organisationales Lernen zu fördern und Fehler und Misserfolge zu entstigmatisieren.⁸⁴

Das Personal des BMZ arbeitet in komplexen und sehr bedeutsamen Themenfeldern. In voller Anerkennung dessen gelingt es der Servicestelle mit dieser Funktion also den Blick auf Evidenz und Wirksamkeit nicht nur im Sinne von optimierter Steuerung und Kontrolle zu richten, sondern auch klar als konstruktive Lernchance für alle zu etablieren.

3.6 Funktion 6: Strategische Beauftragung von Wirkungsevaluierungen

Analog zum SIEL-Modell sollte die Servicestelle die strategische Identifikation von Evidenzbedarfen und die Beauftragung entsprechender Wirkungsmessungen und Wirkungsevaluierungen gemeinsam mit dem Evaluierungsreferat des BMZ koordinieren. Dies beinhaltet die Entwicklung von Learning Agendas für die vier Ziele des BMZ-Reformplans, die zentrale Evidenzlücken identifizieren, und künftige Evaluierungen priorisieren – und die zum Beispiel durch das DEval durchgeführt werden könnten. Ressourcen innerhalb der Programme sollten für Evaluierung zur Verfügung gestellt werden, und ein systematischer Rückfluss der Evaluierungsergebnisse an Entscheidungsträger und Referenten sollte sichergestellt sein.

⁸² Janus, H. & Esser, D. (2025). How Do Foreign Aid Agencies Learn? A Comparative Analysis of Organisational Learning Determinants in National Donor Bureaucracies. [\[Link\]](#)

⁸³ Ibid.

⁸⁴ Allen, S. (2024). Failing and scaling: 4 of BRAC’s latest failures — and biggest learnings. [\[Link\]](#)

3.7 Funktion 7: Vernetzung und internationaler Austausch

Die vorliegende Analyse hat bereits auf eine Vielzahl von internationalen Beispielen und möglichen Partnern in der Arbeit zu mehr Evidenz und Wirksamkeit in der EZ verwiesen. Die Servicestelle sollte den kontinuierlichen Austausch mit vergleichbaren Strukturen anderer Geber pflegen und das BMZ in internationalen Netzwerken wie der Smart Buys Alliance (SBA) vertreten. Die SBA bündelt vorhandene Evidenz zu besonders wirksamen Ansätzen. Staatliche Partner sind derzeit Norad, Sida und das niederländische Außenministerium, weitere Partner zeigen Interesse an einem Beitritt. Eine deutsche Beteiligung könnte die international koordinierte Identifikation und Finanzierung bewährter Ansätze signifikant stärken.

3.8 Funktion 8: Aufbau einer Wirksamkeitsdatenbank für die dt. EZ

In Zusammenarbeit mit internationalen Partnern sollte die Servicestelle an der strategischen und global koordinierten Schaffung von Evidenz zur Wirksamkeit von Entwicklungsprojekten mitwirken. Langfristig könnte sie dabei eine öffentlich zugängliche Plattform und Datenbank bewährter Ansätze aufbauen, die Referenten, Durchführungsorganisationen und Partnerländern als Referenz dient. Das FCDO hat mit seinen sektoralen Best-Buy-Dokumenten, die Interventionen nach Kosten-Effektivität und Evidenzstärke kategorisieren, gezeigt, wie ein solches Instrument die Programmgestaltung verbessern kann. Das Center for Global Development (CGD) hat zu Recht darauf hingewiesen⁸⁵, dass solche Dokumente zu wertvoll sind, um als interne Papiere verborgen zu bleiben: Ihre Veröffentlichung schafft ein öffentliches Gut und stärkt zugleich die Evidenznutzung in Partnerländern.

Ein vergleichbares Referenzprojekt ist auch die IDEAL-Plattform, die aktuell vom Strategic Impact Evaluation Fund der Weltbank aufgebaut wird ([siehe 2.7](#)).

In Deutschland kann das BMZ für eine Wirksamkeitsdatenbank auf ein dichtes Netzwerk deutscher (DEval, IDOS, ISDC, C4ED, GIGA und WZB) und internationaler Forschungspartner (J-PAL, IPA, CEGA, GPRL, CGD und 3ie) zurückgreifen. Zur besonderen Einbeziehung lokalen Wissens und lokaler Wissensproduktion zeigen beispielsweise Projekte wie ACRES in Uganda ([siehe 2.5](#)) und PACE in Südafrika ([siehe 2.6](#)), wie kontextspezifische Evidenz durch Rapid-Evidence-Services im Globalen Süden erzeugt und für Entscheidungstragende aufbereitet werden kann. Die Erfahrung von GIZ und KfW mit der Entwicklung von Standardwirkungsmodellen, wie im Reformplan vorgesehen, kann mittelfristig zur Kartierung von Evidenzlücken (Evidence Gap Maps)⁸⁶ und sektoralen Wirksamkeitslisten beitragen.

3.9 Funktion 9: Strategische Kommunikation von Wirksamkeit

Die Servicestelle sollte in enger Abstimmung mit dem Kommunikationsreferat des BMZ stehen, um die öffentliche Wahrnehmung der Wirksamkeit des BMZ zu beobachten und auf Basis verfügbarer Evidenz strategisch die Wirksamkeit der deutschen EZ gegenüber Parlament und Öffentlichkeit zu kommunizieren. In Zeiten schrumpfender Budgets und

⁸⁵ CGD (2025). FCDO's Best Buys Deserve a Larger Audience. [\[Link\]](#)

⁸⁶ Siehe z.B. 3ie (n.d.). Evidence Gap Maps. [\[Link\]](#)

wachsenden Legitimationsdrucks der EZ ist dies keine Nebenfunktion, sondern eine strategische Notwendigkeit.

Die Scorecard der Weltbank ([siehe 2.7](#)) liefert hier ein einschlägiges Anschauungsbeispiel, wie die aggregierte Wirkung des BMZ im Allgemeinen und die Standardindikatoren im Speziellen mittelfristig nach außen kommuniziert werden können. Auch Deutschland trägt über seine Beiträge an die Weltbank in signifikantem Maße zu den „Scores“ bei, was ihnen aber bisher nur implizit ist. Eine Explizierung dieses Anteils – koordiniert zwischen multilateraler Institution und der Servicestelle – würde umfassender die Wirkungen der deutschen EZ im Mehrebenenansatz kommunizieren.

Das CGD hat gezeigt, dass britische ODA-Ausgaben zehnmal intensiver evaluiert werden als andere Staatsausgaben.⁸⁷ Auf jede Milliarde Pfund EZ-Ausgaben kommen 13 Evaluierungen, verglichen mit nur einer Evaluierung pro Milliarde Pfund sonstiger Staatsausgaben. Auch in Deutschland wird die EZ als das am stärksten evaluierte Politikfeld bezeichnet. Wenn das BMZ die Evidenzbasis dieser Evaluierungen und Wirkungen ausbaut, und diese obendrein strategisch kommuniziert, schafft es eine starke Argumentationsgrundlage in Haushaltsverhandlungen und kann um noch mehr Rückhalt in der Bevölkerung werben.

Exkurs: Evidenzbasierte Steuerung in der Praxis der deutschen EZ

Die deutsche EZ verfügt bereits über Beispiele, die zeigen, wie eine iterativ verstärkte Orientierung an rigoroser Evidenz aussehen kann. Diese Erfahrungen sollten als Ausgangspunkt für die Arbeit der Servicestelle dienen.

Strukturbildende Übergangshilfe (BMZ und GIZ): Das Referat für strukturbildende Übergangshilfe im BMZ macht bereits vor, wie eine schrittweise, verstärkte Orientierung an vorliegender rigoroser Wirkungsevidenz in die operative Arbeit integriert werden kann. In der Vergangenheit wurden gezielte Evidenz-Reviews beauftragt,⁸⁸ fachliche Austauschformate organisiert und das eigene Evidenzverständnis öffentlich kommuniziert.⁸⁹ Das Merkblatt zur Antragstellung für Förderungen im Titel fordert explizit Wirkungsmessungen und die Darlegung externer wissenschaftlicher Evidenz.⁹⁰ Auf Arbeitsebene werden Projektentwürfe gegen Evidenz-Checklisten geprüft.

⁸⁷ CGD (2025b). UK Aid is Ten Times More Scrutinised than Other Government Spending. [\[Link\]](#)

⁸⁸ Siehe statt vieler Murphy, K. et al. (2024). Evidence Review for Food Security, Nutrition, and Climate-Resilience Interventions. Development Innovation Lab, University of Chicago. [\[Link\]](#)

⁸⁹ BMZ, G22 (2025). Forschung für mehr Wirkung in der internationalen Krisenbewältigung. [\[Link\]](#)

⁹⁰ BMZ, Referat KWI (2023). Merkblatt für Zuwendungen des BMZ für die Förderung von Vorhaben des Titels Krisenbewältigung und Wiederaufbau; Infrastruktur, kurz strukturbildende Übergangshilfe des BMZ (KWI). Anlage 2. [\[Link\]](#)

Graduierungsansatz in Somalia (GIZ): Die GIZ implementiert in Somalia den Graduierungsansatz – einen der am besten evaluierten Interventionstypen zur Armutsbekämpfung weltweit – und verbessert ihn durch lokale Adaptation iterativ.⁹¹ Die Servicestelle könnte die herausragende Güteklasse der Evidenz in solchen Projekten herausstellen und sie als eine der Grundlagen für weitere Programmlinien der Armutsbekämpfung verfügbar machen.

Integriertes Wirkungsmanagement (KfW): Die KfW baut seit 2020 ein konzernweites Wirkungsmanagement auf. Ausgehend von der Formulierung einer portfolioübergreifenden und integrierten *Theory of Change*, hat sie KfW-weite Wirkungsindikatoren entwickelt und ihr Indikatorenset strukturiert und offen dargelegt.⁹² Sie bezieht in ihre Steuerung und ihr Reporting sowohl ex-ante als auch ex-post Wirkungsindikatoren mit ein.⁹³ (Insbesondere Ersteres ist gemeinhin in der EZ noch zu selten der Fall.) Auch aktuelle Herausforderungen und Grenzen des Wirkungsmanagements werden klar benannt.⁹⁴

Standardindikatoren und Entwicklung von Standardwirkungsmodellen (BMZ, GIZ und KfW): Das BMZ hat gemeinsam mit GIZ und KfW 43 Standardindikatoren entwickelt und 2022 eingeführt, um fundierter und klarer die aggregierten Wirkungen der deutschen EZ öffentlich kommunizieren zu können.⁹⁵ Die Standardindikatoren sind auch eine erste Grundlage für eine Harmonisierung von Berichtssystemen und für vergleichende Portfolioanalysen. Zugleich reicht ihre bisherige Konzeption noch nicht aus, um portfolioübergreifende Steuerungs- und Skaleneffekte auszulösen. Mit der im Reformplan geforderten Entwicklung von Standardwirkungsmodellen ergibt sich eine wertvolle Ergänzung zur Konsolidierung des aktuellsten Evidenzstands. Zugleich können beide Modelle mittelfristig zur Erstellung von Evidenzkarten (Evidence [Gap] Maps) und sektoralen „What Works“-Listen beitragen, die wiederum die wirkungsorientierte Haushaltssteuerung unterstützen.

All diese Beispiele zeigen: Es gibt im deutschen EZ-System bereits hilfreiche Grundlagen, die den Aufbau einer stärker evidenz- und wirkungsorientierten Steuerung erleichtern können. Die Aufgabe der Servicestelle ist es, diese Ansätze systematisch zusammenzuführen und so zu konfigurieren, dass sie auf das gesamte Portfolio des BMZ und in die Steuerungs-, Mittelallokations-, und Strategieebene ausgeweitet werden können.

⁹¹ GIZ (o. J.). Teilhabe von Frauen am sozialen und wirtschaftlichen Leben in Kismayo fördern. [\[Link\]](#)

⁹² KfW (2025). Das Wirkungsmanagement der KfW. [\[Link\]](#)

⁹³ Ibid.

⁹⁴ Ibid.

⁹⁵ Janus, H. & Esser, D. (2022). Neue Standardindikatoren für die deutsche Entwicklungszusammenarbeit: Was bringen „Zahlen auf Knopfdruck“? IDOS Policy Brief. [\[Link\]](#)

4. Institutionelle Ausgestaltung

Um die oben genannten Funktionen realistisch erfüllen zu können, braucht es eine passende institutionelle Ausgestaltung für die Servicestelle Evidenz. In diesem Abschnitt soll diese durch eine klare Funktionsabgrenzung zum DEval, Vorschläge zur institutionellen Verortung, und Empfehlungen zur Personal- und Mittelausstattung der Servicestelle skizziert werden.

4.1 Klare Funktionsabgrenzung

Eine der zentralen Fragen bei der Einrichtung der Servicestelle ist ihr Verhältnis zum Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval). Die Evidenzarchitektur im Vereinigten Königreich bestehend aus ICAI, K4DD und SIEL bietet hier eine klare Orientierung:

Dimension	DEval (analog ICAI)	Servicestelle (analog K4DD/SIEL)
Funktion	Unabhängige, extern mandatierte Evaluierung	Interne, nachfrageorientierte Evidenznutzung und -beauftragung
Zielgruppe	Parlament, Öffentlichkeit, Fachcommunity	BMZ-Referenten, Leitungsebene, GIZ, KfW, Parlament ad-hoc
Zeitlogik	Retrospektiv (Was hat gewirkt?)	Simultan und prospektiv (Was wirkt aktuell und wird voraussichtlich wirken?)
Unabhängigkeit	Institutionell unabhängig vom BMZ	Eingebettet in BMZ, fachlich unabhängig
Zweck	Rechenschaftslegung	Lernen und Steuerung

DEval führt retrospektive, thematische Evaluierungen durch und berichtet öffentlich. Die Servicestelle hingegen bereitet Evidenz prospektiv für kommende Programmkonzeptionen auf. Sie arbeitet ressortintern – unter anderem auch mit der Leitungsebene zur wirkungsorientierten Mittelallokation – und ist näher an den strategischen und politisch steuernden Gegebenheiten der EZ. Beide Funktionen, DEval und Servicestelle, sind für eine wirksame EZ unverzichtbar, müssen aber institutionell klar getrennt bleiben.

Die Servicestelle sollte DEval-Evaluierungen systematisch für ihre Arbeit nutzen und umgekehrt DEval über die prospektiven Evidenzbedarfe des BMZ informieren.

4.2 Institutionelle Verankerung

Bereits frühere BMZ-Reformprozesse hatten sich das Ziel gesetzt, das BMZ stärker auf Evidenz- und Wirkungsorientierung zu fokussieren. Um jedoch tatsächlich den gewünschten Paradigmenwechsel von Prozess- zu Wirkungslogik einzuleiten, und Evidenz- und Wirkungsperspektiven auch stärker vor die Haushalts- und Programmsteuerung zu bringen, müsste die Servicestelle Evidenz eine klar hervorgehobene Stellung und direkte Berichtslinie an die Leitungsebene haben, etwa in Form eines Stabs auf Ebene einer Abteilungsleitung. Fachlich sollte die Servicestelle im Abteilungsstrang G (Grundsätze) positioniert sein, dort

finden sich auch mehrere einschlägige Referate zur Zusammenarbeit, insbesondere in der Unterabteilung Daten und Wirksamkeit. Die Rolle eines „Leitenden Wirkungsanalysten“ (LWA)⁹⁶ würde der Servicestelle nach innen und außen eine starke Repräsentation geben. Sie könnte, ähnlich wie im Falle der „Chief Economists“ von USAID oder FCDO, durch zeitlich befristete Abordnungen renommierter Wissenschaftler europäischer Universitäten ausgefüllt werden, was frische externe Impulse und eine zusätzliche Fachlichkeit auf dem aktuellen Stand der Forschung einbringen würde.

Sind Evidenz-Einheiten hingegen zu weit von der Entscheidungsebene entfernt, verlieren sie an Einfluss auf die tatsächliche Mittelallokation und es besteht ein erhebliches Risiko, dass sie als praxisferne Zusatzübung wahrgenommen werden.

Das BMZ-Datenlabor, Fachvertretungen von GIZ (Stabstelle Evaluierung, SV Wirksamkeit und Evidenz) und KfW (Abteilung FZ-Evaluierung) sowie Vertretungen von DEval und IDOS wären ständige institutionelle Partner der Servicestelle, mit denen in unterschiedlichen Intensitäten regelmäßige Austauschformate und Schnittstellen etabliert werden sollten. Weitere Akteure sollten auf Einladung hinzukommen können, etwa Universitäten, internationale Evidenzpartner, unabhängige Think Tanks oder die Breite der Zivilgesellschaft.

Eine mögliche institutionelle Verortung und Verankerung ist auch in folgender Abbildung 1 zusammengefasst:

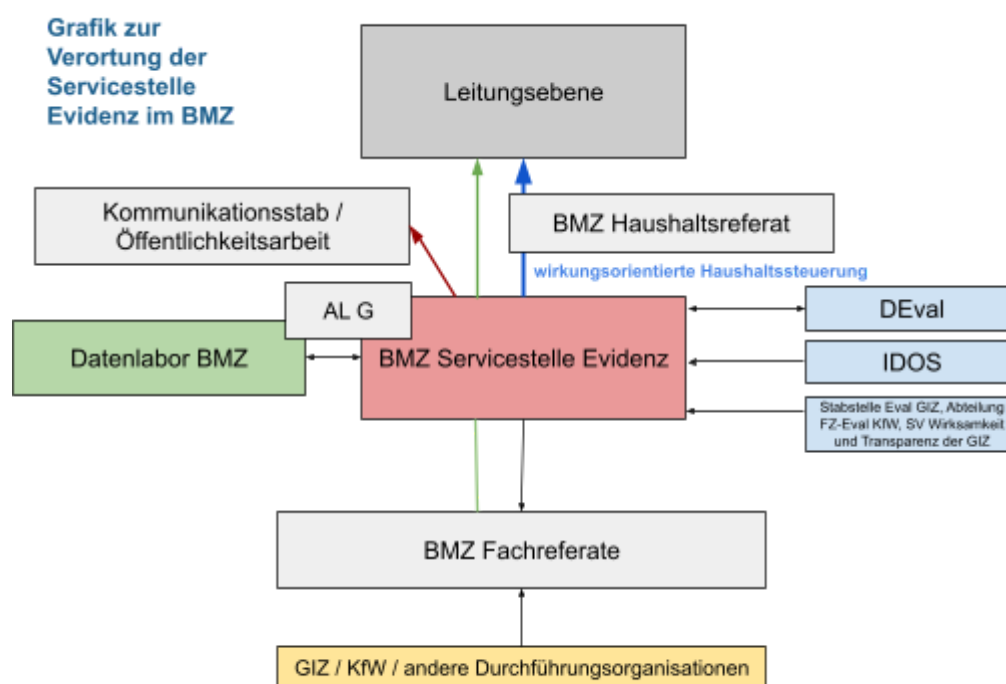


Abb. 1: Grafische Darstellung der institutionellen Einbettung einer Servicestelle Evidenz im BMZ

⁹⁶ Die spezielle Rollenbezeichnung als „Leitender Wirkungsanalyst“ trägt hier dem Verständnis Rechnung, dass die Arbeit zu Evidenz und Wirksamkeit in der EZ kein rein ökonomistisches Unterfangen ist, sondern immer eine mehrdimensionale Perspektive einnimmt.

4.3 Personal und Ressourcenausstattung

Empfohlen wird folgende Zusammensetzung:

- **Leitung:** Mindestens Ebene einer Referatsleitung, bestenfalls Rolle eines „Leitenden Wirkungsanalysten“, mit direktem Berichtskanal an BMZ-Leitung
- **Personal:** 1 VZÄ als Leitung und minimum 6 VZÄ für die Dossiers
 - a. Evidenzsynthese und Portfolioanalyse anhand von Wirkung;
 - b. Kosteneffektivitätsanalysen;
 - c. Datenmanagement;
 - d. Wirksamkeitsberichterstattung;
 - e. Kapazitätsaufbau und Lernkultur.

Angesichts der Komplexität und des Stellenwerts der Aufgaben sollte der Personalumfang mittelfristig auf 8-12 VZÄ anwachsen. (Zum Vergleich: Das USAID OCE verfügte über mehr als 30 Mitarbeitende.)

- **Mittelausstattung:** Eigenes Funktionspostfach, perspektivisch Plattform für Evidenzübersichten und Informationsdiffusion, KI-gestütztes Tool zur Evidenzsynthese, Mittel zur Beauftragung von Programmen zu strategischen Wirkungsevaluierungen und -messungen, sowie unterstützender Studien zu iterativen Wirkungsmessungen und Portfolioanalysen; optional und idealerweise auch für ein Vehikel zur evidenzgestützten Testung innovativer neuer EZ-Projekte.

In Summe und mit Blick auf Projektvolumina der oben genannten internationalen Beispiele ergibt sich ein **kurzfristiger Zielkorridor für die Mittelausstattung der Servicestelle Evidenz von 5-10 Millionen Euro pro Jahr und ein mittelfristiger Korridor von 10-15 Mio. Euro pro Jahr**. Nimmt man den Umfang der bereits bestehenden einschlägigen Titel für Evidenz und Wirkungsorientierung im Kapitel 2305 des Einzelplans 23 („Forschung, Evaluierung und Qualifizierung in der Entwicklungszusammenarbeit“) hinzu – für 2026 waren diese bei ca. 35 Mio. Euro⁹⁷ – ergeben sich Gesamtausgaben für eine strukturierte und zielorientierte Ausstattung zur Stärkung von Wirkung und Evidenz im BMZ von 0,5% des Gesamthaushalts.

Die Kosten-Nutzen-Abschätzung, die die Servicestelle Evidenz für das BMZ-Portfolio schrittweise etablieren soll, lässt sich auch auf sie selbst beziehen: Wenn die Arbeit der Servicestelle Evidenz und ihrer begleitenden Evaluierungsmaßnahmen zu einer Effizienzsteigerung und Kosteneffektivität in nur 1–2 Prozent im Gesamtportfolio des BMZ führte – etwa durch die Umschichtung von Mitteln aus weniger wirksamen in mindestens doppelt so wirksame Programme – entspricht dies einem Wirkungszugewinn von 95–190 Millionen Euro und damit einem geschätzten Return on Invest (ROI) von 6:1 bis 12:1.

Mit Blick auf die Umschichtungen, die das USAID OCE in zweieinhalb Jahren erreichte ([siehe 2.1](#)), ebenfalls nur ausgestattet mit einem Bruchteil des USAID Gesamtbudgets, kann dies als konservative Schätzung betrachtet werden.⁹⁸

⁹⁷ BMF (2025). Bundeshaushaltsplan 2026. Einzelplan 23. [\[Link\]](#)

⁹⁸ Für ein Beispiel zu vergleichbaren und stärkeren Rentabilitäten institutionalisierter Evidenzarbeit, siehe auch Baker, P. and Guzman, J. (2024). Building Institutions for Priority Setting in Health: Six Lessons from 10 Years of iDSI. [\[Link\]](#)

Grundsätzlich ist festzuhalten: Spezifische Investitionen in Evidenz und Wirkungsorientierung gehören zu den rentabelsten Ausgaben im gesamten Einzelplan 23, weil sie Hebelwirkungen für die Kosteneffektivität aller anderen Ausgaben auslösen.

4.3 Erfolgsfaktoren und Risiken

In diesem Abschnitt werden nochmal knapp die Erfolgsfaktoren und mögliche Risiken in der Ausgestaltung der Servicestelle Evidenz benannt.

4.3.1 Erfolgsfaktoren

Politische Rückendeckung und Leadership-Signal: Stärkste Grundvoraussetzung für den Erfolg einer Evidenzeinheit ist die sichtbare Unterstützung durch die Leitungsebene. Das BMZ hat mit dem Reformplan ein ähnliches Signal gesetzt, und steht nun vor dessen Institutionalisierung.

Nachfrageorientierung statt Angebotsorientierung: Evidenzeinheiten scheitern, wenn sie als akademische Parallelstruktur wahrgenommen werden. Die Servicestelle muss von Anfang an auf die konkreten Bedarfe von Referaten und Leitungsebene ausgerichtet sein.

Neue Dynamik, mit Schnittstellen an bestehende Systeme: Die Servicestelle ermöglicht eine neue und moderne Rollenverteilung und Verwaltungsarbeit. Dafür schafft sie Schnittstellen zu bestehenden und bewährten Evaluierungs-, Monitoring- und Wissensmanagementsysteme der deutschen EZ, und entwickelt neue Analysen und Prozesse ohne Doppelstrukturen.

Iterativer Aufbau: Internationale Vorbilder wie K4DD zeigen, dass Evidenzeinheiten und ihre Dienstleistungen über Jahre hinweg iterativ verbessert werden. Die Servicestelle des BMZ sollte mit einem Mindestangebot starten und schrittweise ausgebaut werden.

Transparenz über die eigene Wirkung: Die Servicestelle sollte selbst evidenzbasiert arbeiten und regelmäßig dokumentieren, welchen Einfluss ihre Empfehlungen auf Programmkonzeption und Mittelallokation hatten.

4.3.2 Risiken und Gegenmaßnahmen

Risiko 1: Unzureichendes Mandat: Wenn die Servicestelle lediglich beratend tätig wird, ohne dass ihre Empfehlungen in verbindliche Prozesse einfließen, bleibt ihr Einfluss marginal. Gegenmaßnahme: Verankerung von Mandaten als verbindliche Prozessschritte, z.B. die Kosten-Effektivitäts-Checks für bestimmte Schwellenwerte in der Programmplanung.

Risiko 2: Widerstand etablierter Strukturen und Evidenzsysteme: Andere Einheiten und Organisationen könnten die Servicestelle als zusätzliche Kontrollinstanz wahrnehmen. Gegenmaßnahme: Frühzeitige Einbindung dieser Stellen als Partner. Learning Arenas nach dem Vorbild des Development and Learning Lab ([DLL, 2.4](#)) können Vertrauen aufbauen.

Risiko 3: Überschneidung mit DEval und anderen Einheiten. Unklare Zuständigkeiten könnten zu Doppelarbeit führen. Gegenmaßnahme: Klare formelle Zuständigkeiten durch Geschäftsverteilungsplan und regelmäßige Koordinierungsrunden.

Risiko 4: Unterfinanzierung. Unzureichende Mittel schwächen die Servicestelle in ihrer Lieferfähigkeit. Gegenmaßnahme: Erhöhung des Titels für Forschung und Evaluierung als Teil der Umsetzung des Reformplans.

5. Schlussfolgerung

Mehrere Gebernationen haben sich die letzten Jahre auf den Weg gemacht, eine verstärkte Arbeit zu Evidenz und Wirksamkeit in ihren „ODA-Portfolien“ durch gezielte Einheiten und Programme zu institutionalisieren. Dass nun auch Deutschland – als eine der größten Gebernationen der Welt – durch seinen BMZ-Reformplan eine ähnliche Ambition formuliert, ist sehr erfreulich. Die großen Herausforderungen für die weltweite und auch deutsche EZ – überproportionale Kürzungen; Skepsis der Bevölkerung zu Wirksamkeit bei gleichzeitig immer noch grundsätzlicher Befürwortung; multiple Krisen durch Krieg, Klimawandel und Naturkatastrophen – erhalten so eine wichtige, möglichkeitsorientierte Perspektive. Die Entwicklungszusammenarbeit kann und soll „Deutschlands Rolle in der Welt“ prägen. Als kontinuierlich und wohl am systematischsten evaluiertes Politikfeld kann sie mit der Servicestelle den Beweis antreten, auch weiterhin sowohl im gründlichen Blick auf sich selbst, als auch auf seine Wirkungen in Zielländern, zu lernen, sich anzupassen, und zu innovieren.

Doch jegliche Bemühungen um eine Servicestelle Evidenz werden sich auch ganz entscheidend daran messen lassen müssen, ob sie Evidenz- und Wirkungsperspektiven nicht nur *nach* der Vollendung jeweiliger Programme und Projekte einnehmen (aus einer Logik der Rechenschaftspflicht), sondern wirklich auch *ex-ante* und während der Durchführung (Logik der Wirkungsorientierung). Kooperation Global hat mit diesem Dokument einen Ausgestaltungsvorschlag vorgelegt, der auf die gelebte Praxis in sechs Ländern und zwei internationalen Organisationen blickt, und daraus Empfehlungen für die Servicestelle Evidenz in Deutschland ableitet. Bestenfalls kann dieser Vorschlag als anregende Diskussionsgrundlage dienen.

Quellenverzeichnis

Alle aufgeführten Links wurden zuletzt am 31. Juli 2026 aufgerufen.

- 3ie. (o. J.). *Evidence Gap Maps*. [\[Link\]](#)
- ACRES. (o. J.). *The Evidence to Policy (E2P) Training Programme*. [\[Link\]](#)
- ACRES. (2026a). *Transforming decision-making in Africa*. [\[Link\]](#)
- ACRES. (2026b). *Digital Evidence Hub Dr-RES*. [\[Link\]](#)
- ARI-Database (n.d.). *Areas of Research Interests and Questions of the FCDO*. [\[Link\]](#)
- Allen, S. (2024). *Failing and scaling: 4 of BRAC's latest failures – and biggest learnings*. [\[Link\]](#)
- Baker, P., & Guzman, J. (2024). *Building institutions for priority setting in health: Six lessons from 10 years of iDSI*. [\[Link\]](#)
- BMF. (2025). *Bundeshaushaltsplan 2026. Einzelplan 23*. [\[Link\]](#)
- BMZ. (2026). *Zukunft zusammen global gestalten*. [\[Link\]](#)
- BMZ, G22. (2025). *Forschung für mehr Wirkung in der internationalen Krisenbewältigung*. [\[Link\]](#)
- BMZ, Referat KWI. (2023). *Merkblatt für Zuwendungen des BMZ für die Förderung von Vorhaben des Titels Krisenbewältigung und Wiederaufbau; Infrastruktur (strukturbildende Übergangshilfe des BMZ, KWI) (Anlage 2)*. [\[Link\]](#)
- BRH. (2023). *Bericht über die Evaluierung von Maßnahmen der Entwicklungszusammenarbeit*. [\[Link\]](#)
- CGD. (o. J.). *Project Resource Optimization (PRO)*. [\[Link\]](#)
- CMI. (2026). *CMI awarded funding for new Knowledge Centre to strengthen evidence in aid*. [\[Link\]](#)
- DEval. (o. J.). *Das RIE-Förderprogramm*. [\[Link\]](#)
- DEval. (2026). *Meinungsmonitor Entwicklungspolitik*. [\[Link\]](#)
- DFID. (2011). *DFID's approach to value for money (VfM)*. [\[Link\]](#)
- Dioptra. (o. J.). *Dioptra – Rapid, rigorous cost-efficiency analysis*. [\[Link\]](#)
- Development Learning Lab. (o. J.). *More learning, better outcomes*. [\[Link\]](#)
- Development Learning Lab. (2024). *Annual report 2024* (S. 10). [\[Link\]](#)
- Development Learning Lab. (2025). *Annual report*. [\[Link\]](#)
- FCDO. (2020). *FCDO's review of the Independent Commission for Aid Impact*. [\[Link\]](#)
- FCDO. (2023). *Strategic Impact Evaluation and Learning (SIEL) programme – Call for proposals*. [\[Link\]](#)
- FCDO. (2024). *Chief Economist and Director of Analysis – Candidate brief*. [\[Link\]](#)
- FCDO. (2026). *FCDO programme operating framework*. [\[Link\]](#)
- Feng, X., Kim, D. D., Cohen, J. T., et al. (2020). *Using QALYs versus DALYs to measure cost-effectiveness: How much does it matter?* [\[Link\]](#)

- Frijters, P., Krekel, C., Sanchis, R., & Santini, Z. I. (2024). *The WELLBY: A new measure of social value and progress*. [\[Link\]](#)
- Georgalakis, J. (2026). *Navigating shifting policy priorities in global evidence synthesis*. [\[Link\]](#)
- GIZ. (o. J.). *Teilhabe von Frauen am sozialen und wirtschaftlichen Leben in Kismayo fördern*. [\[Link\]](#)
- Glandon, D., Fishman, S., Tulloch, C., et al. (2023). The state of cost-effectiveness guidance: Ten best resources for CEA in impact evaluations. *Journal of Development Effectiveness*, 15(1), 5–16. [\[Link\]](#)
- Gov.uk. (o. J.). *Chief Economist and Director of Analysis Professor Dennis Novy*. [\[Link\]](#)
- ICAI. (o. J.). *How we work*. [\[Link\]](#)
- Innovations for Poverty Action. (2024). *Strategic Impact Evaluation and Learning (SIEL)*. [\[Link\]](#)
- International Development Committee of the UK Parliament. (2025). *Assessing value, ensuring impact: The FCDO's approach to value for money in Official Development Assistance*. [\[Link\]](#)
- IRC. (o. J. a). *Research and innovation at the International Rescue Committee*. [\[Link\]](#)
- IRC. (o. J. b). *Explore cost analysis at the IRC*. [\[Link\]](#)
- IRC. (o. J. c). *Education Research in Conflict and Protracted Crisis (ERICC)*. [\[Link\]](#)
- IRC. (2023). *Interactive Outcomes and Evidence Framework*. [\[Link\]](#)
- Janus, H., & Esser, D. (2022). *Neue Standardindikatoren für die deutsche Entwicklungszusammenarbeit: Was bringen „Zahlen auf Knopfdruck“?* IDOS Policy Brief. [\[Link\]](#)
- Janus, H., & Esser, D. (2025). *How do foreign aid agencies learn? A comparative analysis*. [\[Link\]](#)
- Jing, W., Tasci, K., Ergun, Ö., Vosti, S., & Webb, P. (2025). *Enhancing cost-efficiency and effectiveness in USAID's food aid supply chain operations in Ethiopia*. [\[Link\]](#)
- K4DD. (2026). *The Knowledge for Development and Diplomacy programme*. [\[Link\]](#)
- KfW. (2025). *Das Wirkungsmanagement der KfW*. [\[Link\]](#)
- Leight, J. (2022). *Patterns of randomization, geographic concentration and funding in development economics*. [\[Link\]](#)
- Lloyd, R., Tjønneland, E., et al. (2024). *Evaluation of Norad's use of knowledge in portfolio management*. [\[Link\]](#)
- Murphy, K., et al. (2024). *Evidence review for food security, nutrition, and climate-resilience interventions*. [\[Link\]](#)
- Norwegian Task Force on evidence-based aid. (2023). *Impact evaluations in Norwegian development cooperation*. [\[Link\]](#)
- PACE. (2026). *Pan-African Collective for Evidence*. [\[Link\]](#)
- REAP. (2026). *Responsive Evidence System for African Policy Needs*. [\[Link\]](#)
- Ritchie, E. (2025). *FCDO's Best Buys deserve a larger audience*. CGD. [\[Link\]](#)
- Ritchie, E. (2026a). *How is the UK aid watchdog viewed within FCDO?* CGD. [\[Link\]](#)

- Ritchie, E. (2026b). *UK aid is ten times more scrutinised than other government spending*. CGD. [\[Link\]](#)
- Ruiz, J. (2025). *How to fix foreign aid – Interview with Dean Karlan*. [\[Link\]](#)
- Smart Buys Alliance. (2026). *The Smart Buys Alliance*. [\[Link\]](#)
- USAID. (2021). *ADS Chapter 201. Program cycle operational policy*. [\[Link\]](#)
- USAID. (2024a). *USAID position paper: Cost-effectiveness*. [\[Link\]](#)
- USAID. (2024b). *Promoting Impact and Learning with Cost-Effectiveness Evidence (PILCEE)*. [\[Link\]](#)
- US-Kongress. (2026). *H.R.7148 – Consolidated Appropriations Act, 2026 (Sec. 7011)*. [\[Link\]](#)
- Weltbank. (o. J. a). *Development Impact Group – Artificial Intelligence*. [\[Link\]](#)
- Weltbank. (o. J. b). *Development Impact Group – Driving policy impact*. [\[Link\]](#)
- Weltbank. (o. J. c). *Dimewiki*. [\[Link\]](#)
- Weltbank. (o. J. d). *Brief – Impact Data and Evidence Aggregation Library (IDEAL)*. [\[Link\]](#)
- Weltbank. (o. J. e). *ImpactAI – Better evidence. Better decisions*. [\[Link\]](#)
- Weltbank. (o. J. f). *Microdata Library – Development Impact Evaluation (DIME)*. [\[Link\]](#)
- Weltbank. (o. J. g). *The World Bank Group Scorecard*. [\[Link\]](#)
- Weltbank. (o. J. h). *The Strategic Impact Evaluation Fund (SIEF)*. [\[Link\]](#)
- Wencker, T. (2022). *Die Verteilung von Mitteln für die deutsche öffentliche Entwicklungszusammenarbeit*. DEval. [\[Link\]](#)
- WHO. (2024). *The art and science of promoting evidence-informed decision-making: A global living evidence map*. [\[Link\]](#)
- Zable, A., & Verhulst, S. (2025). *Monitoring the re-use and impact of non-traditional data*. [\[Link\]](#)

Annex: Umsetzungszeitplan

Der folgende Zeitplan ist infolge ihrer Nennung im Reformplan eine Annäherung an mögliche Zeiträume und Meilensteine für den Aufbau und die Umsetzung der Servicestelle Evidenz im BMZ. Idealerweise kann sie mit einigen ihrer Funktionen bereits zur Programmplanung und Haushaltsaufstellung der nächsten Jahre beitragen.

Phase	Zeitraum	Meilensteine
Phase 1: Konzeption	ca. 2 Monate	• Institutionelle Anbindung in Abteilung G und Berichtslinien festlegen • Personalgewinnung starten (Leitung + 5 Referenten) • Koordinierungsmechanismus mit DEval vereinbaren • Konsultation internationaler Vorbilder (DLL, K4DD, SIEL) • Mandat und Kernfunktionen der Servicestelle formal verabschieden (Geschäftsverteilungsplan)
Phase 2: Aufbau	ca. 6 Monate	• Anknüpfungspunkte in der BMZ Haushalts- und Programmsteuerung (insb. iPA-Prozess) identifizieren und testen • Rapid-Evidence-Helpdesk operationalisieren (Funktionspostfach, Pilotanfragen) • Erste Identifikation bewährter Ansätze mit dem besten Verhältnis von Wirksamkeit, Wirtschaftlichkeit und Umsetzungsfähigkeit; erste Identifikation weniger wirksamer Maßnahmen • Standardisierte Checkliste für Kosten-Effektivitäts-Check entwickeln • Erste Learning Arenas mit GIZ und KfW zu Standardwirkungsmodellen
Phase 3: Integration	ca. 6 Monate	• Kosten-Effektivitäts-Check als Standardschritt in Programmkonzeption verankern • Ersten jährlichen Portfolio-Wirksamkeitsbericht erstellen • Evidenzbasierte Empfehlungen für kommendes Haushaltsjahr erarbeiten • Kapazitätsaufbau: Erste Impact-Evaluation-Workshops für Programmverantwortliche • Digitale Plattform für Helpdesk und Wirksamkeitsdaten im Zusammenspiel mit der Entwicklung eines KI-gestützten Tools zur Evidenzsynthese pilotieren
Phase 4: Konsolidierung	Nach ca. 1,5 Jahren	• Vollständige Integration in die Steuerungsprozesse des BMZ • Personalausbau auf Zielgröße (8–12 Stellen) • Learning Agendas für alle vier BMZ-Reformziele entwickeln • Aufbau Wirksamkeitsdatenbank nach SIEF/IDEAL-Vorbild • Systematische Verankerung internationaler Partnerschaften (SBA, DLL, K4DD)

August 2026

Autoren: Enoch Tabak
Gabriel Hanrieder
Johanna Wicke

Kooperation Global e.V.

Müllerstraße 138D
13353 Berlin
Deutschland

Kontakt: info@kooperationglobal.de
www.kooperationglobal.de



Kooperation Global ist ein gemeinnütziger Think Tank für evidenz- und wirkungsorientierte Entwicklungszusammenarbeit. Wir entwickeln evidenzbasierte Politikvorschläge und unterstützen die Schaffung institutioneller Strukturen für eine wirksamere Entwicklungspolitik.