

Evaluierungspolicy

Standards, Prozesse und strategische Ziele des Evaluierungssystems der Internationalen Zusammenarbeit



Inhaltsverzeichnis

Abkürzungsverzeichnis.....	3
1. Einleitung.....	4
2. Funktion der Evaluierung.....	5
3. Qualitätsstandards.....	6
4. Qualitätssicherung.....	8
Internes Vorgespräch.....	8
Gutachtendenauswahl.....	8
Externes Vorgespräch.....	10
Terms of Reference & Erklärung über organisationale und persönliche Unabhängigkeit.....	11
Inception Report.....	11
Validierungsworkshop.....	12
Annotierte Berichtsvorlage für Langfassung des Evaluierungsberichtes.....	12
Abschlussgespräch.....	12
Lerndokumentation und Follow-up.....	13
Metaevaluierungen.....	13
Qualitätssicherung der Projekte (9.3).....	13
5. Festlegen des Evaluierungsprogrammes.....	13
6. Evaluierungstypen, -ansätze, Erhebungsmethodik und Analysemethoden.....	15
7. Evaluierungskriterien.....	15
8. Capacity WORKS.....	17
9. Kombination von Evaluierung und Projektprüfung.....	17
10. Nutzen und Nutzung von Evaluierungsergebnissen.....	18

Abkürzungsverzeichnis

AGL	Arbeitsgruppenleitung
BMWK	Bundesministeriums für Wirtschaft und Klimaschutz
BMZ	Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung
DAC	Development Assistance Committee
DeGEval	Gesellschaft für Evaluation e.V.
DEval	Deutschen Evaluierungsinstitut der Entwicklungszusammenarbeit
EVA	Evaluierungsverantwortlicher/Evaluierungsverantwortliche
GL	Geschäftsleitung
HG	Hauptgutachtende oder leitende Gutachtende
IR	Inception Report
OECD	Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
PK	Projektkoordinator oder Projektkoordinatorin
PKG	Projektkoordinator oder Projektkoordinatorin im Gutachtendenteam
PTB	Physikalisch-Technische Bundesanstalt
TG	Technische Gutachtende
QI	Qualitätsinfrastruktur
RL	Referatsleitung
RIE	Rigorous Impact Evaluation
ToR	Terms of Reference

1. Einleitung

Diese Evaluierungspolicy erläutert die Standards, Prozesse und strategischen Ziele der Evaluierungspraxis der Internationalen Zusammenarbeit der Physikalisch-Technischen Bundesanstalt (PTB). Die Evaluierungspolicy richtet sich explizit an Projektteams, Gutachtende, Partnerinstitutionen, BMZ-Referate sowie die interessierte Öffentlichkeit.

Die PTB, das nationale Metrologieinstitut, ist eine wissenschaftlich-technische Bundesoberbehörde im Geschäftsbereich des Bundesministeriums für Wirtschaft und Klimaschutz (BMWK). Die PTB ist oberste Instanz bei allen Fragen des richtigen und zuverlässigen Messens und beschäftigt rund 2000 Mitarbeitende. Im Rahmen ihrer Mission unterstützt die PTB, über die Gruppe 9.3 Internationale Zusammenarbeit, die Harmonisierung des internationalen Messwesens, sowie den Auf- und Ausbau einer kompetenten Qualitätsinfrastruktur (QI) in Entwicklungs- und Schwellenländern. Dies trägt wiederum zur nachhaltigen wirtschaftlichen, sozialen und ökologischen Entwicklung bei. Die Projekte der Internationalen Zusammenarbeit werden zumeist im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), aber auch anderer Geber, wie der Europäischen Union, durchgeführt.

Das Evaluierungssystem der PTB richtet sich aus nach den BMZ-Leitlinien „Evaluierung in der Entwicklungszusammenarbeit“ (BMZ, 2021a) und den BMZ-Orientierungsleitlinien „Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit“ (2021b). Die Leitlinien des BMZ wiederum fußen auf den Standards des Entwicklungsausschusses der OECD (Development Assistance Committee, DAC) und den Standards der Gesellschaft für Evaluation (DeGEval). Gemeinsam bilden die Leitlinien und Standards einen Rahmen mit Evaluierungsgrundsätzen und wichtigen Qualitätsstandards, auf welchen die PTB aufbaut und ihr Evaluierungssystem im Rahmen dieser Policy verortet.

Im Einklang mit den genannten Standards und Leitlinien verpflichtet sich die PTB ihre Projekte zu evaluieren und die **Rechenschaftspflicht** gegenüber der Öffentlichkeit und dem Auftraggeber sowie das **organisatorische Lernen** aus gemachten Erfahrungen zu gewährleisten. Von besonderer Bedeutung ist hierfür die Unabhängigkeit der Evaluierungseinheit, welche gemäß der BMZ-Leitlinien (2021a) durch die Trennung vom operativen Projektgeschäft gewährleistet wird. Seit August 2020 gehört die Evaluierung nicht länger der projektkoordinierenden Gruppe 9.3 „Internationale Zusammenarbeit“ an, sondern wurde in die Arbeitsgruppe 9.01 „Prozesse der Internationalen Zusammenarbeit“ integriert. Es sind somit jeweils unterschiedliche Leitungen für das operative und evaluierende Geschäft zuständig. Dies ermöglicht die Koordinierung von Projektevaluierungen von einer zentralen, unabhängigen Stelle. Um die Unabhängigkeit zusätzlich zu stärken, werden externe Gutachtende zur Durchführung der Evaluierungen eingesetzt. Die Auswahl dieser Gutachtenden erfolgt auf der Basis von öffentlich zugänglichen und transparenten Kriterien (Kapitel 4, Gutachtendenauswahl).

In den nächsten Kapiteln werden die Standards und Prozesse der Evaluierungspraxis der PTB sowie strategische Ziele zur Weiterentwicklung des Evaluierungssystems beschrieben. Kapitel 2 erläutert die Funktion von Evaluierung für die Internationale Zusammenarbeit der PTB. In Kapitel 3 werden die Qualitätsstandards der Evaluierungseinheit der PTB dargestellt und Qualitätssicherungsprozessen zugeordnet. Kapitel 4 beschreibt die Qualitätssicherungsprozesse und den Ablauf einer typischen PTB-Evaluierung. Kapitel 5 erläutert den Prozess zur Festlegung des Evaluierungsprogramms. Kapitel 6 geht

auf die angewendeten Evaluierungstypen, -ansätze, -methodik und Analysemethoden ein. Kapitel 7 beschreibt die angewendeten Evaluierungskriterien und deren Operationalisierung. Kapitel 8 erklärt die Rolle von Capacity WORKS im Evaluierungsprozess. Kapitel 9 erläutert die Vorgehensweise bei einer gleichzeitigen Durchführung einer Evaluierung mit einer Projektprüfung. Schließlich beschreibt Kapitel 10 die Prozesse zur Erhöhung des Nutzens und der Nutzung von Evaluierungsergebnissen.

2. Funktion der Evaluierung

In Übereinstimmung mit den Leitlinien (BMZ, 2021a) werden Evaluierungen in der PTB definiert als: „Systematische und objektive Analysen und Bewertungen von laufenden oder abgeschlossenen Entwicklungsmaßnahmen. Diese Untersuchungen umfassen in der Regel Konzeption, Umsetzung und insbesondere Ergebnisse der Entwicklungsmaßnahmen und sollen handlungsrelevante Erkenntnisse und in geeigneten Fällen Empfehlungen für Verbesserungen der Gestaltung enthalten (nach OECD DAC 1991 bzw. 2002/2009).“

Zusätzlich zu dieser Definition legt die PTB gemäß den Leitlinien (BMZ, 2021a) zwei Funktionen fest, welche Evaluierungen erfüllen sollen: 1) Die Schaffung von Transparenz und Erbringung der Rechenschaftspflicht, sowie 2) Wissensmanagement und organisationales Lernen.

Transparenz und Rechenschaftspflicht sollen gegenüber allen Stakeholdern von PTB-Projekten erbracht werden. Hierzu zählen: Das BMZ, die Partnerinstitutionen und die Bürgerinnen und Bürger, welche mit ihren Steuern PTB-Projekte finanzieren. Die Evaluierung schafft Transparenz und Rechenschaftspflicht durch die unabhängige Überprüfung der Projekte bezüglich erzielter Wirkungen, der Nachhaltigkeit dieser Wirkungen und der Effizienz, mit welcher die Wirkungen erreicht wurden. Die Ergebnisse der Evaluierung werden mit dem BMZ und den Partnerinstitutionen geteilt und sind öffentlich auf der Website der PTB¹ einsehbar. Von einer Publikation wird nur in begründeten Einzelfällen abgesehen, wenn Schaden durch die Veröffentlichung der Evaluierung für die an der Evaluierung beteiligten Personen entstehen könnte. Für alle ab 2023 begonnenen Evaluierungen wird außerdem nicht nur der *One-Pager* und eine Kurzfassung, sondern auch die Langfassung des Evaluierungsberichts veröffentlicht.

Das **Wissensmanagement und organisationales Lernen** soll durch Evaluierungen gestärkt werden bei: Partnerinstitutionen, auf der Leitungsebene der projektkoordinierenden Gruppe 9.3, in den Projektteams und in der Evaluierungseinheit. Mit den Partnerinstitutionen finden Workshops zur Validierung der Ergebnisse und zur Diskussion der Empfehlungen statt. Die Empfehlungen richten sich ebenfalls an das Projektteam, die Partner, die Leitungsebene der Gruppe 9.3 sowie an die Evaluierungseinheit. Zudem werden die entstehenden Wissensprodukte mit den Partnerinstitutionen geteilt und die Umsetzung der Empfehlungen vom Projektteam dokumentiert. Innerhalb der PTB finden Präsentationen für die projektkoordinierende Gruppe 9.3 statt, passende Evaluierungen werden an entsprechende Arbeitskreise weitergeleitet und *lessons learned* werden in einem Peer-Austausch zwischen Projektkoordinatoren und Projektkoordinatorinnen aufgearbeitet. Die Umsetzung von PTB-internen Empfehlungen an das Projektteam, die Leitungsebene der Gruppe 9.3 und die Evaluierungseinheit (9.01) werden in Form

¹ Link zur Website eines der Regionalreferates (siehe unten „Evaluierungen“):
<https://www.ptb.de/cms/de/ptb/fachabteilungen/abt9/gruppe-93/ref-931.html>

eines *Follow-Up-Sheets* und einer Lerndokumentation nachgehalten (mehr unter Kapitel 10 Nutzung von Evaluierungsergebnissen).

3. Qualitätsstandards

Wie bereits in der Einleitung der Evaluierungspolicy erläutert, orientiert sich die PTB an den „Qualitätsstandards für die Entwicklungsevaluierung“ der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD, 2010), und den „Standards für Evaluation“ der Gesellschaft für Evaluation e.V. (DeGEval, 2016). Die genannten Standards bilden außerdem die Grundlage für die Leitlinien des BMZ „Evaluierung in der Entwicklungszusammenarbeit“ (BMZ, 2021a) und den Orientierungsleitlinien des BMZ „Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit“ (BMZ, 2021b).

Um darzustellen wie die OECD/DAC und DeGEval Standards sowie deren Einbettung in die BMZ-Leitlinien im PTB-Evaluierungssystem reflektiert sind, müssen die Qualitätsstandards operationalisiert und in greifbare Qualitätskriterien umgewandelt werden. Hierfür nutzt die PTB die Ableitung von Qualitätskriterien aus der „Meta-Evaluierung zur Qualität von (Projekt-)Evaluierungen in der deutschen EZ“ des DEval (Guffler et al., 2022). Diese Qualitätskriterien wurden von der PTB organisationsspezifisch angepasst, erweitert und in der unten dargestellten Tabelle je einem qualitätssichernden Prozess zugeordnet. Die Prozesse gewährleisten die Einhaltung der Qualitätskriterien in den Evaluierungen der PTB und sind in Kapitel 4 „Qualitätssicherung“ ausführlich dargestellt und reflektiert.

Methoden und Berichtslegung

Qualitätskriterium (QK)	Qualitätssicherung (erläutert in Kapitel 4)
1. Beschreibung des Evaluierungsgegenstandes 2. Beschreibung des Kontexts der Entwicklungsmaßnahme 3. Berücksichtigung des Kontexts bei Ergebnissen 4. Beschreibung des Erkenntnisinteresses 5. Angemessenheit des methodischen Vorgehens	Internes und externes Vorgespräch, Inception Report, Annotierte Berichtsvorlage
6. Nachvollziehbarkeit der Informationsquellen 7. Informationsgehalt der Zusammenfassung 8. Anwendung OECD/DAC-Kriterien 9. Kohärenz von Daten-Ergebnissen-Schlussfolgerungen	Annotierte Berichtsvorlage
10. Darstellung der Wirkungszusammenhänge	Annotierte Berichtsvorlage, (Qualitätssicherung der Projekte in 9.3 e.g. Indikatoren im Modulvorschlag)
11. Informationsgehalt Leistungsbeschreibung	Terms of Reference (ToR)

Partizipation, Unabhängigkeit und Fairness

12. Einbindung der Partnerinstitutionen	Internes Vorgespräch, Inception Report, Validierungsworkshop
13. Einbindung der Stakeholder	Internes Vorgespräch, Validierungsworkshop, Abschlusspräsentation
14. Berücksichtigung partnerschaftlicher Ansätze (Beachtung nationaler Evaluierungspolitiken: Policies, nationale Richtlinien, Handreichungen, Evaluierungspläne etc.)	Inception Report
15. Evaluierungsethik (Schutz der Evaluierungsteilnehmenden, Umgang mit sensiblen Daten, Datenschutz)	Internes Vorgespräch
16. Evaluierungseffizienz	
17. Berücksichtigung Kapazitätsentwicklung (durch Einbindung lokaler Kapazitäten)	Gutachtendenauswahl, Internes Vorgespräch
18. Berücksichtigung von Gemeinschaftsevaluierungen	
19. Einbezug von lokalen Gutachtenden	
20. Ausgewogenheit von Männern und Frauen	
21. Kompetenz der Gutachtenden	Gutachtendenpool, Kriterien für Gutachtendenauswahl
22. Organisationale Unabhängigkeit der Gutachtenden	Kriterien für Gutachtendenauswahl, Erklärung über organisationale und persönliche Unabhängigkeit
23. Unvoreingenommenheit der Gutachtenden	
24. Zugänglichkeit des Evaluierungsberichtes (inklusive Zusammenfassung)	Abschlussgespräch
25. Zugänglichkeit für Stakeholder	
26. Transparenz von Meinungsverschiedenheiten zwischen Gutachtenden und PTB	

Nutzen und Nutzung

27. Rechtzeitigkeit der Ergebnisse	Internes und externes Vorgespräch, ToR
28. Evaluierungseffizienz	
29. Ausreichende humane und finanzielle Ressourcen für Evaluierung	
30. Adressatenorientierung der Empfehlungen	Internes und externes Vorgespräch, annotierte Berichtsvorlage, Lerndokumentation und Follow-up
31. Nützlichkeit und Handlungsorientierung der Empfehlungen	
32. Management Response	Jahresgespräch mit 9.3 Leitungsebene

4. Qualitätssicherung

Die Evaluierungseinheit nutzt mehrere Prozesse der Qualitätssicherung, um die Einhaltung der soeben aufgeführten Qualitätskriterien zu gewährleisten:

Internes Vorgespräch

Das interne Vorgespräch wird mit dem Evaluierungsteam, dem Projektteam und der Referatsleitung geführt, während das externe Vorgespräch (s.u.) zusätzlich die jeweiligen Hauptgutachtenden und technischen Gutachtenden umfasst. Als erster Schritt im internen Vorgespräch wird sichergestellt, dass das Projekt ausreichende finanzielle Mittel zur Verfügung hat. Anschließend wird im Sinne der Evaluierungsethik sorgfältig abgewogen, ob auf eine Evaluierung, oder Teile einer Evaluierung, wegen möglicher Risiken oder Schäden für Teilnehmende, verzichtet werden muss. Ist dies nicht der Fall, so beginnt die Planung der Evaluierung. Es werden der grobe Zeitplan der Evaluierung und die geschätzten finanziellen und personellen Ressourcen erörtert (QK 27-29). Mit Bezug auf die Evaluierungseffizienz wird die Möglichkeit einer Remote-Evaluierung oder hybriden Evaluierung geprüft und das vorläufige Erkenntnisinteresse gegen potenzielle Kosten abgewogen. Bei der zeitlichen Planung der Evaluierung wird sichergestellt (wie auch im Evaluierungsprogramm), dass die Ergebnisse in die Prüfung des (potenziellen) Folgeprojektes einfließen können. Sollte kein Folgevorhaben geplant sein, so kann das spezifische Erkenntnisinteresse der Evaluierung besonders auf sektorspezifische (z.B. Energie, Lebensmittel, Textil) oder instrumenten- und/oder prozessspezifische Inhalte gelegt werden. Dies soll ermöglichen die Evaluierungsergebnisse für andere Projekte zu nutzen (QK 30). Abschließend werden unter Anwendung der Kriterien für die Gutachtendenauswahl entsprechende Personalien erörtert.

Gutachtendenauswahl

Die PTB beauftragt unabhängige (externe) Gutachtende für die Durchführung ihrer Evaluierungen. Die Evaluierung ist zwingend von mindestens einer/-m unabhängigen, federführenden Gutachtenden durchzuführen. In der Regel setzt sich das Gutachtendenteam jedoch aus einem/r Hauptgutachtenden (HG), einem/r technischen Gutachtenden (TG) und/oder ggf. einem „Projektkoordinator oder einer Projektkoordinatorin im Gutachtendenteam (PKG)“ zusammen. Alle Mitglieder des Gutachtendenteams nehmen eine aktive Rolle in der Datenerhebung, -analyse, Auswertung und Präsentation der Evaluierungsergebnisse ein. Die genaue Verteilung der Aufgaben ist evaluierungsspezifisch und wird im Vorgespräch erörtert und in den Terms of References festgehalten.

Zur Zusammensetzung des Gutachtendenteams müssen mehrere **Kriterien** beachtet werden:

1. HG oder TG haben mindestens eine Evaluierung in der PTB durchgeführt.
2. Die Gutachtenden dürfen weder 1) politisch, 2) noch operativ, 3) noch beratend an Projekten beteiligt sein und 4) nicht zu den direkten Begünstigten des Projektes gehören, um die Unvoreingenommenheit zu gewährleisten. Daher unterzeichnen die Gutachtenden eine „Erklärung über organisationale und persönliche Unabhängigkeit“, in welcher sie ihre organisationale Unabhängigkeit und persönliche Unvoreingenommenheit schriftlich bestätigen.
3. Es sollte pro Evaluierung mindestens eine lokale/ein lokaler HG oder TG an der Evaluierung teilnehmen. Auch hier gilt es die organisationale und persönliche Unabhängigkeit zu wahren, welche

im kleinen Feld der Qualitätsinfrastruktur oft nur durch den Einsatz eines HG oder TG aus einem anderen Land der Region gewährleistet werden kann.

4. Wenn möglich sollte mit vorhandenen Strukturen (Evaluierungseinheiten, wissenschaftliche Institute o.ä.) des Partnerlandes kooperiert werden, um Evaluierungen gemeinschaftlich durchzuführen.
5. Bei Erfüllung aller vorherigen Kriterien und gleicher Kompetenz, soll auf eine Ausgewogenheit von Männern und Frauen im Gutachtendenteam geachtet werden.

Die folgenden **Kompetenzen** sollten vom Gutachtendenteam abgedeckt werden:

1. Fundierte Evaluierungserfahrung (HG).
2. Fachkompetenz im für das Projekt relevanten Themenbereich wie z.B. Umwelt, Energie, Pharmazie, Wirtschaftsentwicklung etc. (TG, möglichst auch HG).
3. QI-Kenntnisse und -Erfahrung (TG, möglichst auch HG).
4. Landeskenntnisse/Kenntnisse der Region (TG, möglichst auch HG).
5. Entwicklungspolitische Kenntnisse und Erfahrungen, sicherer Umgang mit Verfahren der deutschen Technischen Zusammenarbeit (HG).

Darüber hinaus sollten folgende **allgemeine Anforderungen** erfüllt sein:

- Sensibilität für Querschnittsfragen (Gender, Armut, Soziokultur, Umwelt)
- Sehr gute mündliche und schriftliche Sprachkenntnisse der jeweiligen Evaluierungssprache
- Methodenkompetenz: Angemessener Umgang und Kenntnisse von quantitativen und qualitativen Methoden der empirischen Sozialforschung
- Überdurchschnittlich gute Analysekompetenz, verbunden mit der Fähigkeit, die Analyse und ihre Ergebnisse angemessen in Schriftform zu dokumentieren
- Belastbarkeit, Teamfähigkeit, interkulturell angemessenes Verhalten
- Tropentauglichkeit
- Hauptgutachter oder Hauptgutachterin mit Ausbildungshintergrund in Ökonomie, Natur-, Ingenieur-, Politik- oder Gesellschaftswissenschaften

Gutachtende, welche die oben genannten Kriterien erfüllen, können in den **Gutachtendenpool** aufgenommen werden. Eine Aufnahme erfolgt nach fachlicher Prüfung der Kriterien durch zwei Mitglieder des Evaluierungsteams sowie der Zustimmung der Arbeitsgruppenleitung 9.01.

Die Auswahl von Gutachtenden für eine **Projektauvaluierung** erfolgt in drei Schritten:

1. Im internen Vorgespräch werden aufgrund des Evaluierungsgegenstandes relevante Kompetenzen erörtert, welche für die Gutachtendenauswahl von besonderer Bedeutung sind (beispielsweise Kenntnisse im Fachbereich der Metrologie oder Vertrautheit mit internationalen Wirtschaftsorganisationen).
2. Die Projektseite schlägt daraufhin technische Gutachtende für die Evaluierung vor, während die Hauptgutachtenden von dem/der jeweiligen Evaluierungsverantwortlichen vorgeschlagen werden.
3. Diese Vorschläge werden gemeinsam diskutiert, die finale Entscheidung über die Auswahl der Hauptgutachtenden erfolgt durch den/die Evaluierungsverantwortlichen. Bei der Auswahl des

technischen Gutachtenden obliegt dem/der Evaluierungsverantwortlichen ein Vetorecht, um die Einhaltung der o.g. Kriterien sicherzustellen.

Eine Ausnahme zur Bedingung der operativen Unabhängigkeit für Gutachtende bilden „Projektkoordinierende im Gutachtendenteam“ (PKGs). PKGs werden eingesetzt, um die Lernfunktion von Evaluierungen zu stärken. Ihre Teilnahme an Evaluierungen soll für eigene Lernerfahrungen und Kompetenzaufbau genutzt werden, die dann in der eigenen Projektimplementierung Anwendung finden und dem verstärkten Austausch von *lessons learned* und *best practices* zwischen den Regionalreferaten dienen.

Für die **Teilnahme von PKGs** gelten allerdings einige Kriterien:

- PKGs bekommen eine temporäre Entlastung von den regulären projektbezogenen Aufgaben
- PKGs können keine Evaluierungen des eigenen Regionalreferats begleiten
- PKGs sollten mindestens über eine 2-jährige Berufserfahrung bei der projektkoordinierenden Gruppe 9.3 verfügen
- PKGs müssen die Arbeitssprache des Projektes beherrschen
- PKGs müssen über Erfahrungen in den thematischen Schwerpunkten des zu evaluierenden Projektes verfügen (z.B. durch die Arbeit in den eigenen Projekten)
- Haben mehrere PKs ihr Interesse an einer Evaluierung bekundet, findet die Auswahl aufgrund folgender Kriterien statt: Fachlicher Hintergrund der/des PKG, Interessen für das eigene Projekt und Bezug zu aktuellen oder geplanten Aufgaben.

Die Teilnahme von PKGs an Evaluierungen kann zu Lasten der organisationalen und operativen Unabhängigkeit und somit Glaubwürdigkeit der Evaluierung gehen, da Mitarbeitende des operativen Bereiches ihre Kollegen und Kolleginnen aus anderen Referaten evaluieren und hierdurch potenziell ein Interessenskonflikt entstehen kann. Dieses potenzielle Risiko wird jedoch durch die Zusammenarbeit mit externen HG und TG kompensiert. Außerdem kann eine transparente Darstellung der Beteiligung des/r PKG im Evaluierungsbericht die Glaubwürdigkeit des Evaluierungsberichtes erhöhen. Vorteile der Beteiligung von PKGs sind außerdem eine starke Steigerung der Nützlichkeit, da die Lernerfahrung des/der Projektkoordinierenden direkt auf das eigene Projekt übertragen werden kann, und die Multiplikatorfunktion, welche der/die Projektkoordinierende für das generierte Wissen einnimmt. Aufgrund der positiven Effekte für den Nutzen und die Nutzung von Evaluierungen, wird die potenziell geschwächte – wenn auch kompensierte – Unabhängigkeit und Glaubwürdigkeit in Kauf genommen.

Externes Vorgespräch

Im externen Vorgespräch wird ein gemeinsames Verständnis für den Ablauf der Evaluierung zwischen den Gutachtenden, dem Projektteam, den Evaluierungskordinierenden und der Referatsleitung geschaffen. Besonders im Fokus steht hier die Nützlichkeit, welche durch die Einbindung der relevanten Stakeholder sowie der Erörterung ihres Erkenntnisinteresses und ihrer speziellen Evaluierungsfragen gewährleistet wird (QK 1-4, 12, 13). Das Projektteam wird im internen Vorgespräch dazu aufgefordert die Bedarfe und Erkenntnisinteressen der Partnerinstitutionen in Steuerungstreffen des Projektes zu sammeln und aktiv einzubringen. Für einen effizienten Verlauf der Evaluierung wird zudem eine klare Rollenverteilung der beteiligten Parteien festgehalten sowie ein klares Verständnis für die annotierte

Berichtsvorlage und den Inception Report (QK 1-4) geschaffen. Basierend auf dem Erkenntnisinteresse schlagen die Gutachtenden das Evaluierungsdesign vor, welches diskutiert und bei Bedarf angepasst wird (QK 5). Das externe Vorgespräch kann fallabhängig zweigeteilt werden, um den Gutachtenden mehr Zeit zu geben ein bedarfsgerechtes Evaluierungsdesign zur Diskussion zu stellen. Für den Fall, dass sich eine Remote-Evaluierung oder Hybrid-Evaluierung anbietet (bspw. bei Global- oder Regionalprojekten), wird auf den PTB-Leitfaden: *Remote-Evaluierungen – erste Erfahrungen und Empfehlungen* (Lange, Telfser & Wallerath, 2020) verwiesen. Dieser enthält Empfehlungen, Lernerfahrungen und Tipps, an welchen sich die Gutachtenden orientieren können.

Terms of Reference und Erklärung über Organisationale und Persönliche Unabhängigkeit

Auf Basis des externen Vorgesprächs werden die Verträge für die Gutachtenden aufgesetzt, welche die ToRs enthalten. In den ToRs und/oder ergänzenden Vertragsbedingungen müssen folgende Punkte aufgegriffen werden: 1. Zweck der Evaluierung, 2. Nutzende, 3. Ziele, 4. Methoden, 5. Finanzielle Mittel (Mengengerüst), 6. Zeitlicher Rahmen, 7. Veröffentlichungsrechte, 8. Aufgabenverteilung und Verantwortlichkeiten von HG und TG. Zusätzlich zu den ToRs müssen die Gutachtenden eine Erklärung über ihre Unabhängigkeit unterschreiben. In der „Erklärung über organisationale und persönliche Unabhängigkeit“ versichern die Gutachtenden, dass diese weder 1) politisch, 2) noch operativ, 3) noch beratend am Projekt beteiligt sind, noch 4) zur Zielgruppe des Projektes gehören sowie organisational unabhängig sind und persönlich unvoreingenommen.

In den ToRs verpflichten sich die Gutachtenden außerdem der Evaluierungsethik der PTB, welche sich am Standard des Roten Kreuzes (IFRC, 2011) orientiert und den DeGEval Standards (DeGEval, 2016), OECD/DAC Standards (OECD, 2010) sowie den BMZ-Leitlinien (2021a) entspricht:

Evaluierungen müssen unter Einbezug von ethischen und rechtlichen Gesichtspunkten durchgeführt werden, wobei das Wohlergehen der an der Evaluierung beteiligten und von ihr betroffenen Personen besonders zu berücksichtigen ist. Evaluierungen sollten sich an internationale sowie nationale Normen und Vorschriften halten, um Risiken, Schäden und Belastungen für die Evaluierungsteilnehmenden so gering wie möglich zu halten. Die Gutachtenden sollten die Bräuche, die Kultur und die Würde der Teilnehmenden respektieren. Dazu gehören Unterschiede aufgrund von Religion, Geschlecht, Behinderung, Alter, sexueller Orientierung und ethnischer Zugehörigkeit. Besondere Aufmerksamkeit sollte der Frage der Diskriminierung und der Ungleichheit der Geschlechter gewidmet werden (in Übereinstimmung mit der Allgemeinen Erklärung der Menschenrechte der Vereinten Nationen). Ebenso gilt der Grundsatz "Do no harm". Gutachtende sollten die Evaluierungsteilnehmenden über ihre Rechte aufklären, ihre Zustimmung einholen, die Vertraulichkeit der Teilnehmenden sicherstellen und jegliche illegale oder potenziell schädliche Aktivität unterlassen.

Inception Report

Der Inception Report dient der Sicherstellung der Evaluierbarkeit und ggf. der Anpassung der Fragestellung sowie der Prüfung der praktischen Durchführbarkeit und Prüfung der Einhaltung von Qualitätsstandards und Querschnittsthemen der BMZ-Leitlinien (2021a): Menschenrechtliche Sorgfaltspflicht,

Chancengleichheit (u.a. Gendersensibilität, *Leave No One Behind*²), Klimasensibilität, Konfliktsensibilität und „Do no harm“. Zudem werden hier verschiedene Methoden- und Berichtslegungsaspekte qualitätsgesichert: Evaluierungsgegenstand, Kontext der Evaluierung, Rahmenbedingungen des Projektes, zentrale Fragestellungen und Erkenntnisinteresse sowie die Datenerhebungsmethoden und Auswertung. Im Inception Report wird außerdem die Auseinandersetzung der Gutachtenden mit partnerschaftlichen Ansätzen (nationale Evaluierungspolitiken, wie Policies, nationale Richtlinien, Handreichungen, Evaluierungspläne etc.) und deren Beachtung im Evaluierungsdesign gefordert. Falls Evaluierungspläne des Partnerlandes ähnliche Evaluierungen vorsehen, wird eine mögliche Zusammenarbeit geprüft und, wenn möglich, in die Wege geleitet. Schließlich wird der Inception Report mit den Partnerinstitutionen zur Kommentierung geteilt. Gibt es keine weiteren ungeklärten Punkte bzgl. des Inception Reports, so wird dieser offiziell abgenommen und die Feldphase bzw. Datenerhebung beginnt.

Validierungsworkshop

Nach Abschluss der Datenerhebungsphase wird ein Validierungsworkshop mit den Partnerinstitutionen, den Gutachtenden, dem Projektteam und ggf. Evaluierungskoordinierten organisiert. Im Workshop wird ein offener und kritischer Austausch angestrebt, bei welchem die Ergebnisse der Evaluierung und die daraus resultierenden Empfehlungen diskutiert werden. Der Workshop dient somit der Einbindung der Partnerinstitutionen (QK 12) sowie dem Nutzen und der Nutzung der Evaluierungsergebnisse.

Annotierte Berichtsvorlage für die Langfassung des Evaluierungsberichtes

Die Langfassung des Evaluierungsberichtes wird kommentiert durch: Die Referatsleitung, das Projektteam, die/den Evaluierungskoordinierten und die Partnerinstitutionen. Anschließend werden die Kommentare von den Gutachtenden angenommen oder mit einer Begründung abgelehnt. Die Kommentierung des Evaluierungsberichtes dient der Qualitätssicherung der im Inception Report dargelegten Anforderungen. Zudem wird geprüft, ob zukünftig eine Evaluierungsmatrix zum Einsatz kommen soll, um die herangezogenen Quellen für bestimmte Berichtsteile und deren Belastbarkeit übersichtlich darzustellen (QK 32). Außerdem geprüft wird der Informationsgehalt der Zusammenfassung, die Anwendung der OECD/DAC Kriterien, die Kohärenz von empirischen Schlussfolgerungen, die Einhaltung der BMZ-Orientierungsleitlinien einschließlich ihrer Evaluierungsfragen sowie die Darstellung der Wirkungszusammenhänge (QK 33-35). Abschließend wird geprüft, dass alle Empfehlungen an einen klaren Adressaten gerichtet sind und für diesen nützliche, praktisch implementierbare Informationen enthalten (QK 30). Der Bericht wird daraufhin zur finalen Richtigstellung eventueller Fehlinformationen an die Partnerinstitutionen zur Durchsicht geschickt (QK 12).

Abschlussgespräch

Das Abschlussgespräch schließt das Projektteam, die Gutachtenden sowie die Evaluierungskoordinierten und die Referatsleitung ein (QK 13). Es dient der Klärung von offenen Fragen, der Diskussion über Empfehlungen und dem Planen der Veröffentlichung des One-Pagers, der Kurzfassung und der Langfassung (QK 5, 23). One-Pager, Kurzfassung und Langfassung des Evaluierungsberichtes werden

² Leave No One Behind (LNOB) oder auf Deutsch „niemanden zurücklassen“, ist eines der Prinzipien der Agenda 2030. Es fordert, dass die Menschen zuerst erreicht werden, die die geringsten Entwicklungschancen haben. Im Kontext des Inception Reports der Evaluierung soll das Prinzip den Anspruch verdeutlichen, dass die Meinungen und Bedürfnisse von vulnerablen und marginalisierten Gruppen in der Evaluierung berücksichtigt werden und auch die Entwicklungsmaßnahme selbst auf ihre Inklusivität untersucht wird.

mit den Partnerinstitutionen geteilt (QK 6, 24). Eventuelle Meinungsverschiedenheiten werden im Abschlussgespräch diskutiert und, falls die Meinungsverschiedenheiten fortbestehen, die Meinung der Gutachtenden in den Evaluierungsbericht übernommen (QK 25). Die PTB behält sich vor, in Einzelfällen eine Stellungnahme an den Evaluierungsbericht anzuhängen (QK 7).

Lerndokumentation und Follow-up

Empfehlungen, welche sich an das Projektteam richten, werden in einem sogenannten „Follow-Up-Sheet“ festgehalten. Das Projektteam dokumentiert fortlaufend die Umsetzung der Empfehlungen oder begründet die Nichtumsetzung im Follow-Up-Sheet. Alle Empfehlungen, welche den operativen Bereich als Ganzes betreffen, werden gesammelt und mit den Referatsleitungen und der Gruppenleitung validiert und diskutiert. Diese geben daraufhin eine Management Response ab, in welcher sie Empfehlungen übernehmen oder begründet ablehnen. Derselbe Prozess erfolgt ebenfalls für die Evaluierungseinheit, welche alle Empfehlungen, die an sie gerichtet sind, sammelt und eine Stellungnahme abgibt.

Metaevaluierungen

Zusätzlich zu den üblichen Qualitätssicherungsmechanismen nimmt die Evaluierungseinheit der PTB auch an Metaevaluierungen teil. Diese erlauben eine externe Prüfung der Evaluierungseinheit und dienen der Identifizierung von Weiterentwicklungspotenzialen. Die Beauftragung einer solchen Evaluierung erfolgt in der Regel durch die PTB, kann aber beispielsweise auch als Teil von DEval Evaluierungen erfolgen.

Qualitätssicherung der Projekte (9.3)

Abgesehen von den Prozessen der Evaluierungseinheit (9.01), haben auch die qualitätssichernden Prozesse der projektkoordinierenden Gruppe 9.3 einen Einfluss auf die Qualität der Evaluierungen. Besonders die Sicherstellung der Verwendung von SMART-Indikatoren³ in den Modulvorschlägen ist essenziell, um Wirkungszusammenhänge plausibel prüfen zu können. Des Weiteren sollen Baseline-Daten, Monitoring-Daten sowie Fortschrittsberichte für die Evaluierung genutzt werden. Aus diesen Gründen soll ein reger Austausch mit den für die Qualitätssicherung und Monitoring zuständigen Personen der projektkoordinierenden Gruppe 9.3 gepflegt werden.

5. Festlegen des Evaluierungsprogrammes

Das jährliche Evaluierungsprogramm legt fest, welche Projekte evaluiert werden sollen. Das Verfahren zur Auswahl der Projekte strebt eine Balance zwischen Rechenschaftspflicht und organisationalem Lernen an, und erfolgt in mehreren Schritten:

- (1) Die Erstellung des Programmes beginnt mit der Erfassung aller Projekte, die im Folgejahr enden. Diese Grundgesamtheit umfasst den Zeitraum Q2 des Folgejahres (e.g. 01.04.2025) bis Ende Q1 des darauffolgenden Jahres (e.g. 31.03.2026).

³ Indikatoren sollten die folgenden Attribute aufweisen: Specific, Measurable, Attainable, Relevant, Timed.

- (2) Die projektkoordinierende Gruppe 9.3 „Internationale Zusammenarbeit“ nominiert bis zu drei Projekte aus der Grundgesamtheit, für welche aus ihrer Sicht ein besonderes Lernpotenzial besteht. Diese Projekte werden in jedem Fall in das Evaluierungsprogramm aufgenommen.
- (3) Es wird eine zufällige Stichprobe von Projekten gezogen, welche einschließlich der von 9.3 nominierten Projekte, in Summe 70% der Grundgesamtheit abdecken.
- (4) Der Jahresplan wird bei Bedarf aktualisiert, beispielsweise im Falle von Projektverlängerungen.

Sollten bestimmte Projekte aus dem Evaluierungsprogramm (bereits gezogene Projekte) unerwartet verlängert werden, so werden diese verpflichtend in das Programm des Folgejahres übernommen, ohne dabei erneut am Losverfahren teilzunehmen. Solche Projekte werden also zusätzlich zu den 70 Prozent der zufällig gezogenen Projekte evaluiert. Sollte es allerdings absehbar sein, dass ein Projekt verlängert wird, so wird dieses nicht in die Grundgesamtheit aufgenommen, sondern nimmt an der Stichprobenziehung des nächsten Jahres teil. Sollte es wider Erwarten dann doch zu keiner Verlängerung kommen, so muss das Projekt noch im selben Jahr evaluiert werden, damit die Stichprobe nicht umgangen werden kann.

Das Verfahren wird dem Anspruch der Rechenschaftspflicht gerecht, indem es eine repräsentative Anzahl von Projekten zufällig auswählt. Außerdem stellt es sicher, dass Projekte mit besonderem Lernpotenzial durch die projektkoordinierende Gruppe 9.3 nominiert werden können und somit nicht durch Zufall aus der Stichprobe fallen. Hierdurch wird auch dem Anspruch des organisationalen Lernens ebenfalls Genüge getan.

ERSTELLUNG EVALUIERUNGSPROGRAMM



Quelle: Eigene Darstellung

6. Evaluierungstypen, -ansätze, Erhebungsmethodik und Analysemethoden

Gemäß den Leitlinien des BMZ (2021a) bedient sich die PTB vorwiegend Schlussevaluierungen, da sich direkte entwicklungspolitische Wirkungen meist frühestens gegen Ende des Projektes einschätzen und evaluieren lassen. Schlussevaluierungen werden zumeist im letzten Halbjahr der Projektlaufzeit durchgeführt und bis Laufzeitende abgeschlossen. Nach Bedarf können zusätzliche Evaluierungen durchgeführt werden, wie Ex-Post-Evaluierungen zur Einschätzung der Dauerhaftigkeit von Wirkungen nach mehreren Jahren, Zwischenevaluierungen oder Querschnittsevaluierungen. In der Zukunft sollen diese weiteren Evaluierungstypen im Portfolio der Evaluierungseinheit einen größeren Platz einnehmen.

Die Evaluierungsmethodik wird an den Evaluierungsgegenstand angepasst und kann sich von Evaluierung zu Evaluierung unterscheiden. Typischerweise sind die Evaluierungen der PTB jedoch theoriebasierte *Realist Evaluations*. Diese gehen von einer nicht deterministischen, sondern probabilistischen Ursachen-Wirkung-Beziehung aus. Die *Realist Evaluation* nimmt außerdem an, dass der Erfolg einer Projektmaßnahme vom Kontext, in welcher die Maßnahme durchgeführt wird, abhängt sowie von den Akteuren, welche sich in diesem Kontext bewegen und ihn auf unterschiedliche Art und Weise wahrnehmen. Als häufigstes Mittel der Datenerhebung werden deshalb Interviews eingesetzt. Online-Umfragen und die Erstellung von deskriptiven Statistiken sowie zusätzliche Erhebungsmethoden können je nach Evaluierungsgegenstand auch zum Einsatz kommen.

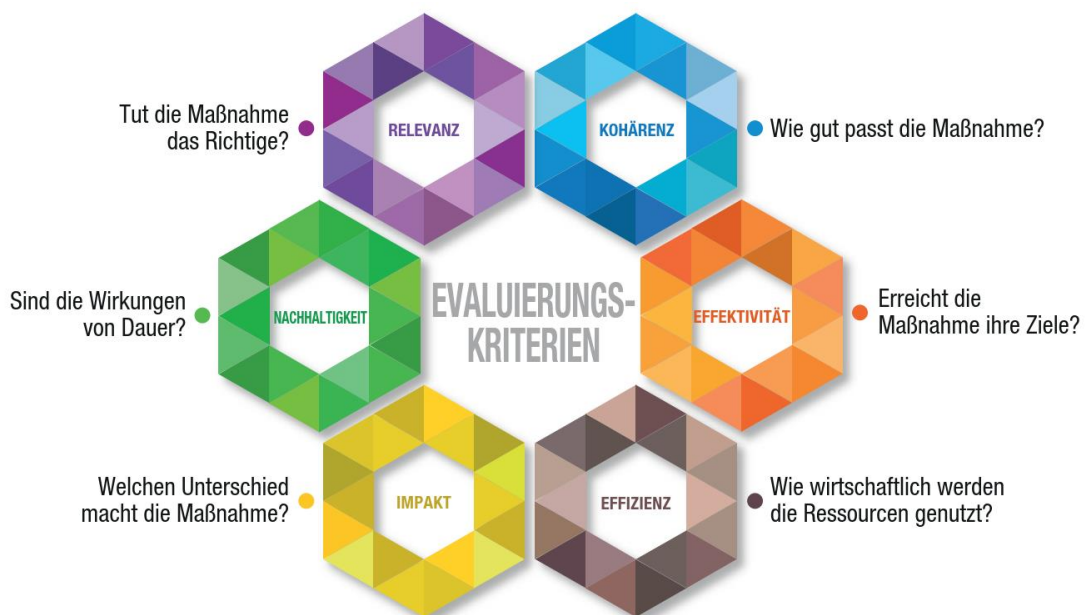
In der Zukunft soll die Methodenvielfalt in Evaluierung der PTB ausgeweitet werden. Hierzu wird in einem kontinuierlichen Prozess die systematische Verankerung von Evaluierungstypen, Evaluierungsansätzen, Erhebungsmethoden und Analysemethoden geprüft. Unter anderem soll in diesem Zusammenhang der kontributionsanalytische Ansatz pilotiert werden, um eine stärkere Fokussierung auf eine Ursachen-Wirkung-Beziehung der Projektlogik zu erreichen. In jeder Evaluierung soll jedoch weiter gegenstandsangemessen entschieden werden, welche Methodik am geeignetsten erscheint und zum Einsatz kommen soll.

Aufgrund der relativ geringen Projektvolumina der PTB im Vergleich zu anderen Durchführungsorganisationen, wie der GIZ oder der KfW, werden *Rigorous Impact Evaluations* (RIEs) vorerst nicht in das Methodenportfolio der PTB aufgenommen. Die Kosten einer solchen Methodik würden nicht im Verhältnis zum Nutzen der Erkenntnisse und des Projektauftragswertes stehen. Somit wäre dem DeGEval-Qualitätskriterium *Evaluierungseffizienz* nicht genüge getan. Die Möglichkeit rigorose Methodiken gewinnbringend einzusetzen, wird jedoch kontinuierlich geprüft und einzelne Pilotierungen werden bei passenden Projekten durchgeführt.

7. Evaluierungskriterien

Die Bewertungen der Entwicklungswirksamkeit werden anhand der sechs OECD/DAC-Kriterien (2010) – Relevanz, Kohärenz, Effektivität, Effizienz, entwicklungspolitische Wirkungen (Impact) und Nachhaltigkeit – und der damit verbundenen Leitfragen vorgenommen. Die weitergehende Operationalisierung der Kriterien und Erstellung von Unterfragen erfolgt entsprechend der BMZ-Orientierungsleitlinien „Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit“ (BMZ, 2021b). Im Sinne des Nutzens

und der Nutzung von Evaluierungsergebnissen können zusätzliche Fragen, die dem Erkenntnisinteresse der Partnerinstitutionen oder des Projektteams dienen, aufgenommen werden.



Quelle: BMZ (2021b)

Diese OECD/DAC-Kriterien werden entlang einer 6-stufigen Notenskala (s. unten) bewertet. Die Gesamtbewertung des Projektes wird anhand des Mittelwertes der Noten für die OECD/DAC-Kriterien gebildet, wobei alle Kriterien gleich gewichtet werden. Falls Effektivität, Impact oder Nachhaltigkeit jedoch mit „4“ oder schlechter bewertet werden, so gilt dies als K.O.-Kriterium und die Gesamtbewertung wird auf „4“ herabgestuft.

Notenskala für die Bewertung der OECD/DAC-Kriterien

Bewertung	Note	Beschreibung
sehr erfolgreich	1	sehr gutes, deutlich über den Erwartungen liegendes Ergebnis
erfolgreich	2	gutes, voll den Erwartungen entsprechendes Ergebnis
eingeschränkt erfolgreich	3	zufriedenstellendes Ergebnis; liegt unter den Erwartungen, aber es dominieren die positiven Ergebnisse
eher nicht erfolgreich	4	nicht zufriedenstellendes Ergebnis; liegt unter den Erwartungen und es dominieren trotz erkennbarer positiver Ergebnisse die negativen Ergebnisse
überwiegend nicht erfolgreich	5	trotz einiger positiver Teilergebnisse dominieren die negativen Ergebnisse deutlich.
gänzlich erfolglos	6	das Projekt ist gänzlich erfolglos, die Situation ist eher verschlechtert

8. Capacity WORKS

Capacity WORKS ist ein Managementmodell zur Projektsteuerung, welches 2006 in der damaligen GTZ entstand. Das Modell umfasst fünf wesentliche Erfolgsfaktoren (Strategie, Kooperation, Steuerungskultur, Prozesse, Lernen und Innovation) und eine Reihe von *Tools* zur Erreichung der Faktoren. Die Anwendung bestimmter Capacity WORKS Tools ist in der PTB verpflichtend. Aufgrund der hohen Bedeutung, die Capacity WORKS in der Durchführung der Projekte einnimmt, wird es auch in den Evaluierungen mitbetrachtet und im Rahmen der Capacity WORKS Selbsteinschätzung bewertet. Die Ergebnisse werden unter dem OECD/DAC Kriterium Effektivität in der Unterfrage zur Durchführungsqualität zusammengefasst.

9. Kombination von Evaluierung und Projektprüfung

Evaluierungen und Prüfungen von Folgevorhaben können in der PTB seit Mitte 2016 kombiniert werden. Evaluierungen müssen nach den Standards der DeGEval (2016) nützlich sein (Nützlichkeit), realistisch und diplomatisch durchgeführt werden (Durchführbarkeit), rechtlich und ethisch korrekt ablaufen (Korrektheit) sowie fachlich angemessene Informationen hervorbringen (Genauigkeit). Die Einhaltung dieser Standards kann durch die Nähe von Evaluierungen zu Prüfungen von Folgevorhaben beeinträchtigt werden. Um dies zu vermeiden, wurde von der Evaluierungseinheit eine Handlungshilfe erarbeitet, die den Prozess an potenziell konfliktären Punkten regelt und eine klare Anleitung für das kombinierte Verfahren liefert.

Bezüglich der Gutachtendenauswahl ist festgelegt, dass Prüfung und Evaluierung vom selben Gutachtenteam durchgeführt werden können, sollte eine ausreichende fachliche und methodische Qualifikation für Prüfung und Evaluierung vorhanden sein. Sollte das Gutachterteam für fachliche Fragen der Prüfung nicht geeignet sein, ist es dem Regionalreferat freigestellt, eine für Prüfungsfragen zuständige Person dem Team hinzuzufügen. Im Vorgespräch soll explizit auf mögliche Konflikte im Spannungsfeld von Prüfung und Evaluierung sowie deren Vermeidung eingegangen werden. Dafür wird die Vorlage für das Vorgespräch um den Punkt „mögliche Konflikte im Spannungsfeld von Prüfung und Evaluierung“ ergänzt. Prüfung und Evaluierung können im Anschluss an das Vorgespräch in ein und demselben Vertrag mit den Gutachtenden geregelt werden, es müssen jedoch separate ToRs genutzt werden. Außerdem wird der Inception Report im Falle eines kombinierten Verfahrens um einen Kurzbericht „Prüfung“ ergänzt. Die Feldphase muss in der Umsetzung zwischen Evaluierung und Prüfung unterscheiden. Treffen mit Partnerinstitutionen können entweder an zwei verschiedenen Terminen stattfinden oder es muss aus dem Interviewleitfaden eine klare Trennung beider Bereiche ersichtlich sein. In diesem Fall muss die zweifache Funktion des Interviews den Interviewpartnern zu Beginn mitgeteilt werden. Außerdem ist der Übergang vom Evaluierungs- zum Prüfungsteil des Interviewleitfadens im Gespräch deutlich zu machen. Es sollte mit dem Evaluierungsteil begonnen werden, um eine mögliche Voreingenommenheit bei den Antworten zum Evaluierungsteil zu verringern, welche durch die Konzentration auf ein Folgevorhaben entstehen könnte. Generell können weitere, im Rahmen der Prüfung anwesende Teilnehmende an kombinierten Gesprächen teilnehmen. Gutachtende haben jederzeit die Möglichkeit den Personenkreis zu reduzieren, bestimmte Teilnehmende erst später zum Gespräch zu bitten oder separate Gespräche zu vereinbaren. Sollte es im Vorfeld für nötig erachtet werden, können zwei sukzessive

Termine für Evaluierung und Prüfung angesetzt werden, um den Teilnehmendenkreis auf die jeweilige Thematik, insbesondere bei voraussichtlich konfliktbeladenen Situationen, zu beschränken. Validierungsworkshop für Prüfung und Evaluierung werden zeitlich getrennt. Außerdem sind für Prüfung und Evaluierung selbstverständlich zwei separate Berichte zu verfassen.

10. Nutzen und Nutzung von Evaluierungsergebnissen

Der Nutzen von Evaluierungsergebnissen besteht im Wesentlichen in der Erbringung der Rechenschaftspflicht gegenüber der Öffentlichkeit und dem BMZ sowie dem organisationalen Lernen.

Zur Erbringung der *Rechenschaftspflicht* gegenüber dem BMZ werden alle Evaluierungsprodukte (*One-Pager*, Kurzzusammenfassung und Langfassung des Evaluierungsberichts) dem Ministerium vorgelegt. Außerdem besteht das Angebot einer Präsentation der Evaluierungsergebnisse für relevante Referate des Ministeriums sowie die Möglichkeit, ein Gespräch zu terminieren. Für die Öffentlichkeit waren bisher nur der *One-Pager* und die Kurzzusammenfassung auf der Website der PTB einsehbar. In Zukunft soll jedoch auch die Langfassung veröffentlicht werden, um eine größere Transparenz gegenüber der Öffentlichkeit zu erzielen.

Um das *organisationale Lernen* möglichst effektiv zu gestalten, sind spezifische Akteurinnen und Akteure in spezifische Informationsformate eingebunden. Für das Projektteam, die Leitungsebene der projektkoordinierenden Gruppe 9.3 und die Evaluierungseinheit werden Follow-Up-Sheets und Lerndokumentationen verwendet, um die Umsetzung von Empfehlungen nachzuhalten. Die Empfehlungen müssen entweder angenommen oder begründet abgelehnt werden. Die *Follow-Up-Sheets* erfüllen somit nicht nur eine Kontrollfunktion, sondern bieten auch eine Form der Dokumentation, die eine systematische Analyse der Umsetzung von Empfehlungen ermöglicht. Die Ergebnisse von PTB-Evaluierungen und Meta-Evaluierungen, an welchen die PTB sich beteiligt, werden an fachspezifische Arbeitskreise der projektkoordinierenden Gruppe 9.3 weitergeleitet.

Die Ergebnisse werden ebenfalls in Form einer Präsentation mit der Gruppe 9.3 geteilt, wozu auch PTB-externe Personen eingeladen werden, wie zum Beispiel Gutachtende aus dem Gutachtendenpool und in den PTB-Projekten eingesetzte Kurzzeitexperten. Die interessierte Öffentlichkeit, die Wissenschaft und andere Durchführungsorganisationen haben die Möglichkeit, die veröffentlichten Evaluierungsberichte über die Website der PTB zu nutzen. Hierzu soll in Zukunft eine Datenbank mit Filterfunktion für eine einfachere Zugänglichkeit entstehen. Schließlich wird das Lernen auf der Seite der Partnerinstitutionen durch eine aktive Einbindung in den Evaluierungsprozess, einen Workshop zur Validierung und Diskussion der Ergebnisse sowie durch das Teilen der Evaluierungsprodukte facilitiert. Zukünftig sollen weitere Formate zur Vergemeinschaftung von in der Evaluierung identifizierten guten Praktiken und *lessons learned* entwickelt werden. Diese Formate sollen in Wissensmanagement- und Einarbeitungsprozessen verankert werden, um die Nutzung von Evaluierungsergebnissen auch bei Personalfluktuation zu gewährleisten.

Literaturverzeichnis

- Lange, S., Telfser, K., & Wallerath, S. (2020). *Remote-Evaluierungen erste Erfahrungen und Empfehlungen*. Ptb.de. https://www.covid19.ptb.de/fileadmin/covid19/downloads/PTB_Info_Remote_Evaluieren_DE.pdf
- Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. (2021a). *BMZ-Papier 4. Leitlinien für die Evaluierung der Entwicklungszusammenarbeit*. Bmz.de. <https://www.bmz.de/resource/blob/92884/bmz-leitlinien-evaluierung-2021.pdf>
- Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. (2021b). *Grundsätze der Evaluierung. Evaluierungskriterien der bilateralen Entwicklungszusammenarbeit. Orientierungslinien des BMZ*. Bmz.de. <https://www.bmz.de/resource/blob/92894/evaluierungskriterien.pdf>
- DeGEval-Gesellschaft für Evaluation e. V. (2016). *Standards für Evaluation*. Degeval.org. https://www.degeval.org/fileadmin/Publikationen/DeGEval-Standards_fuer_Evaluation.pdf
- Evaluierung*. (o. J.). Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. Abgerufen 7. Februar 2023, von <https://www.bmz.de/de/ministerium/evaluierung>
- Guffler, K., L. Kunert, M. Wittenberg und N. Herforth (2022), *Meta-Evaluierung zur Qualität von (Projekt-)Evaluierungen in der deutschen Entwicklungszusammenarbeit*, Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Bonn.
- International Federation of Red Cross and Red Crescent Societies. (2011). *IFRC Framework for Evaluation*. Ifrc.org. <https://www.ifrc.org/sites/default/files/2021-09/IFRC-Framework-for-Evaluation.pdf>
- OECD. (2010). *DAC-Reihe Leitlinien und Grundsatztexte: Qualitätsstandards für die Entwicklungsevaluierung*. OECD. <https://doi.org/10.1787/9789264085183-de>