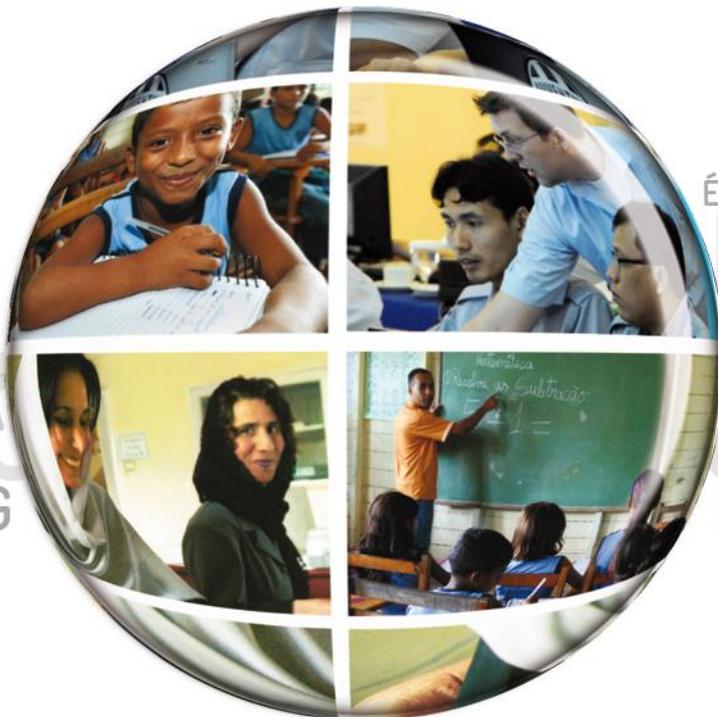


Stabsstelle Monitoring und Evaluierung



Querschnittsauswertung Bildung: Meta-Evaluierung und Synthese

Impressum

Als Bundesunternehmen unterstützt die GIZ die deutsche Bundesregierung bei der Erreichung ihrer Ziele in der internationalen Zusammenarbeit für nachhaltige Entwicklung.
Bei der Darstellung handelt es sich um ehemalige GTZ/InWEnt-Aktivitäten, die bedingt durch die Umfirmierung als GIZ-Aktivitäten bezeichnet werden.

Herausgeber

Deutsche Gesellschaft für
Internationale Zusammenarbeit (GIZ) GmbH

Friedrich-Ebert-Allee 40
53113 Bonn, Deutschland
T +49 228 44 60-1877
F +49 228 44 60-2877

Dag-Hammarskjöld-Weg 1-5
65760 Eschborn, Deutschland
T +49 61 96 79-14 08
F +49 61 96 79-80 14 08

E evaluierung@giz.de
I www.giz.de/monitoring

Verantwortlich

Martina Vahlhaus

Autoren

Dieses Gutachten wurde von unabhängigen externen Sachverständigen erstellt.
Es gibt ausschließlich deren Meinung und Wertung wieder.

Prof. Dr. Stephan Gerhard Huber
dipl.päd. Eveline Steinger
Dr. Pierre Tulowitzki
BA IR Marco Wenger
lic.oec.publ. Marcus Büzberger

Institut für Bildungsmanagement und Bildungsökonomie (IBB) & Institut für Internationale Zusammenarbeit (IZB)
Pädagogische Hochschule Zug
Zugerbergstrasse 3
6301 Zug, Schweiz

Erscheinungsort und -jahr

Bonn, Eschborn, Zug 2014

Inhalt

Inhalt	3
Abkürzungsverzeichnis	5
Zusammenfassung	7
1 Einleitung	14
1.1 Übersicht	14
1.2 Datengrundlage	15
2 Metaevaluierung	17
2.1 Methodik der Metaevaluierung	17
2.2 Ergebnisse	20
2.3 Aspekte der Auswertung und Vergleichbarkeit	22
3 Synthese	23
3.1 Methodik der Synthese	23
3.2 Ergebnisse	25
3.2.1 Manageriale Erfolgs- und Misserfolgskriterien	25
3.2.1.1 Unabhängige Evaluierungen	25
3.2.1.2 Projektfortschrittskontrollen	29
3.2.1.3 Dezentrale Evaluierungen	32
3.2.2 Sektorspezifische bzw. fachliche Erfolgs- und Misserfolgskriterien	33
3.2.2.1 Unabhängige Evaluierungen	33
3.2.2.2 Projektfortschrittskontrollen	38
3.2.2.3 Dezentrale Evaluierungen	40
3.2.3 Querschnittsthemen	41
3.2.3.1 Unabhängige Evaluierungen	41
3.2.3.2 Projektfortschrittskontrollen	41
3.2.3.3 Dezentrale Evaluierungen	42
3.2.4 Übergreifende Fragen	43
3.2.4.1 Unabhängige Evaluierungen	43
3.2.4.2 Projektfortschrittskontrollen	48
3.2.4.3 Dezentrale Evaluierungen	51
4 Fazit	53
5 Empfehlungen	58
6 Anhang	61
6.1 Ampelsystem	61
6.2 Übersicht der zu untersuchenden Evaluierungen	66
6.3 Evaluierungsfragen der ToR	68
6.4 Analyseraster Meta-Evaluierung	70
6.5 Analyseraster Synthese	74
6.6 Benotung der DAC-Kriterien und der Beurteilung der Angemessenheit der Benotung	76
6.7 Übersicht über Programmenden in den einzelnen Vorhaben der UEs	77
6.8 Tabellen zu den Aspekten der Vergleichbarkeit (UE, PFK, DE)	78
6.9 Tabellen zur Wirkungsmessung auf Zielgruppen- und Mittelebene	81
6.10 Übersicht der Empfehlungen und der Basis für Empfehlungen	89

6.11	Datenblätter (Meta-Evaluierung)	92
6.11.1	Analyseraster Meta-Evaluierung UE Ghana	92
6.11.2	Analyseraster Meta-Evaluierung UE Guinea	96
6.11.3	Analyseraster Meta-Evaluierung UE Honduras, Guatemala, Peru	100
6.11.4	Analyseraster Meta-Evaluierung UE Indonesien	104
6.11.5	Analyseraster Meta-Evaluierung UE Namibia	108
6.11.6	Analyseraster Meta-Evaluierung UE Pakistan	112
6.11.7	Analyseraster Meta-Evaluierung UE Peru	116
6.11.8	Analyseraster Meta-Evaluierung UE Tschad	120
6.11.9	Analyseraster Meta-Evaluierung UE Zentralamerika	124
6.11.10	Analyseraster Meta-Evaluierung UE SV Innovation	128
6.11.11	Analyseraster Meta-Evaluierung PFK Afghanistan	132
6.11.12	Analyseraster Meta-Evaluierung PFK Benin	136
6.11.13	Analyseraster Meta-Evaluierung PFK Guinea	140
6.11.14	Analyseraster Meta-Evaluierung PFK Honduras	144
6.11.15	Analyseraster Meta-Evaluierung PFK Kongo	148
6.11.16	Analyseraster Meta-Evaluierung PFK Kosovo	152
6.11.17	Analyseraster Meta-Evaluierung PFK Malawi	156
6.11.18	Analyseraster Meta-Evaluierung PFK Mauretanien	160
6.11.19	Analyseraster Meta-Evaluierung PFK Mosambik	164
6.11.20	Analyseraster Meta-Evaluierung PFK Sri Lanka	168
6.11.21	Analyseraster Meta-Evaluierung PFK überregional (Sektorvorhaben)	172
6.11.22	Analyseraster Meta-Evaluierung PFK Zentralasien	176
6.11.23	Analyseraster Meta-Evaluierung DE Kosovo	180
6.11.24	Analyseraster Meta-Evaluierung DE Zentralamerika	184
6.12	Literatur	188

Abkürzungsverzeichnis

A	Beurteilung der Angemessenheit der Benotung (durch Verfasser dieser Querschnittsanalyse)
Abb.	Abbildung
ADEA	Association for the Development of Education in Africa (Gesellschaft für Bildungsentwicklung in Afrika)
AFG	Afghanistan
ANG	angemessen
Anz	Anzahl der Beurteilungen
B	Benotung der DAC-Kriterien durch Evaluatoren
BEN	Benin
BMZ	Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung
CHA	Tschad
COD	Kongo
DAC	Development Assistance Committee
DE	Dezentrale Evaluierung
EduPAZ	Educación para la Paz
EFA	Education for All – Fast Track Initiative (Bildung für Alle – Schnellspur-Initiative)
Erg.	Ergebnis/Ergebnisse
EZ	Entwicklungszusammenarbeit
FIERE	Filles éduquées réussissent (gebildete Mädchen sind erfolgreich)
FZ	Finanzielle Zusammenarbeit
GHA	Ghana
GPE	Global Partnership for Education (Globale Partnerschaft für Bildung)
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit
GUI	Guinea
HGP	Honduras, Guatemala, Peru
HON	Honduras
IBB	Institut für Bildungsmanagement und Bildungsökonomie
INA	Indonesien
InWEnt	Internationale Weiterbildung und Entwicklung
IZB	Institut für Internationale Zusammenarbeit in Bildungsfragen
KOS	Kosovo
M+E-System	Monitoring- und Evaluierungssystem
MAW	Malawi
MDG	Millenium Development Goals
MOZ	Mosambik
MTN	Mauretanien
NAM	Namibia
NRO	Nichtregierungsorganisation, nichtstaatliche Organisation
OECD	Organisation for Economic Co-operation and Development
PAK	Pakistan
PASEC	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN

PER	Peru
PFK	Projektfortschrittskontrolle
QSA	Querschnittsauswertung
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SRI	Sri Lanka
SV	Sektorvorhaben
ToR	Terms of Reference
TZ	Technische Zusammenarbeit
UE	Unabhängige Evaluierung
UNDP	United Nations Development Programme
VG	Vergleichsgruppe
ZA	Zentralamerika
ZAM	Dezentrale Evaluierung Zentralamerika
ZAs	Zentralasien
ZG	Zielgruppe
ZM	zu mild
ZS	zu streng

Zusammenfassung

Das Institut für internationale Zusammenarbeit in Bildungsfragen (IZB) und das Institut für Bildungsmanagement und Bildungsökonomie (IBB) der Pädagogischen Hochschule Zug verfassten im Auftrag der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH eine Querschnittsauswertung (QSA) bestehend aus einer Meta-Evaluierung und einer Synthese von Evaluierungen von Projekten der GIZ im Sektor Bildung. Ziel der Querschnittsauswertung war eine zusammenfassende Analyse der Einzelevaluierungen, um Ergebnisse zu aggregieren, Erfolgs- und Misserfolgskriterien zu identifizieren und Handlungsempfehlungen abzuleiten.

Datengrundlage bilden zehn unabhängige Evaluierungen (UEs), welche die Stabsstelle Monitoring und Evaluierung hat durchführen lassen, zwölf Projektfortschrittskontrollen (PFKs) von Bildungsvorhaben sowie zwei dezentrale Evaluierungen (DEs). Für die Meta-Evaluierung wurde ein Bewertungssystem entwickelt, welches zunächst Fragen nach der Qualität der einzelnen Evaluierungen beantwortet, um Evaluierungsberichte, die als nicht ausreichend robust eingeschätzt werden, aus der Datengrundlage der Synthese ausschließen zu können.

Die Ergebnisse der Meta-Evaluierung der UEs fallen positiv aus: In allen Bewertungskriterien werden mindestens ausreichende Ergebnisse erzielt. Hervorzuheben sind die methodische Transparenz, die verwendeten Datenbasen sowie die Empfehlungen. Im Durchschnitt wurden UEs von uns mit 15.7 von 18 möglichen Punkten bewertet, wobei der Minimalwert bei sechs Punkten liegt. Bezüglich der PFKs zeigt sich ein gemischteres Bild: Hier beträgt die von uns vergebene Durchschnittsbewertung 12 von 18 möglichen Punkten. Drei der untersuchten PFKs erreichten nicht die Voraussetzungen, um als robust eingestuft zu werden. PFK-übergreifend lässt sich festhalten, dass im Bereich der angemessenen Verwendung von Wirkungsketten¹ sowie im Bereich der methodischen Transparenz Spielräume der Verbesserung existieren. Auf Ebene der DEs erhielt eine DE 13, die andere 12 von 18 möglichen Punkten.

Die untersuchten UEs und PFKs wurden von den Evaluatoren im Durchschnitt anhand der Kriterien des Ausschusses für Entwicklungshilfe (Development Assistance Committee – DAC) der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) in puncto Relevanz mit 1.6, in puncto Effektivität mit 2.4, in puncto Impact mit 2.9, in puncto Effizienz mit 2.7 und in puncto Nachhaltigkeit mit 2.9 bewertet. Der Durchschnitt der von den Evaluatoren vorgenommenen Gesamtbewertungen betrug 2.7, was ein Ergebnis zwischen Stufe 2 (gutes Ergebnis, ohne wesentliche Mängel) und Stufe 3 (zufriedenstellendes Ergebnis; es dominieren die positiven Ergebnisse) darstellt. Aus unserer Sicht wurde in den uns vorgelegten Evaluierungen tendenziell ein wenig zu milde bewertet. Als methodisches Grundgerüst für die Erstellung der Synthese wurde ein detailliertes Analyseraster verwendet, welches aus den Fragen und Anliegen in den Reference (ToR) abgeleitet wurde. Die Daten wurden für die drei Evaluierungstypen (UE, PFK, DE) jeweils separat ausgewertet. Dabei wurden die Evaluierungsberichte inhaltsanalytisch ausgewertet und wiederkehrende Punkte identifiziert, die drei übergreifenden Themenfeldern zugeordnet werden können: Erfolgsfaktor Kooperation, Monitoring und Evaluierung (im Bildungssektor), und fachliche Aspekte im Bildungssektor.

¹ Der Begriff „Wirkungskette“ wurde in der GIZ zunächst durch den Begriff „Wirkungsgefüge“ und schließlich vom Begriff „Wirkungsmodell“ abgelöst. Da in den evaluierten Dokumenten jedoch überwiegend von „Wirkungsketten“ gesprochen wird, greifen wir in diesem Text auf diese Formulierung zurück.

Erfolgsfaktor Kooperation

Im Bereich der managerialen Erfolgs- und Misserfolgskriterien war die **Gestaltung der Kooperationsbeziehungen** in den evaluierten Projekten ein wiederkehrendes Thema. Bemängelt wurde die Kooperation hinsichtlich Motivation und Kapazitäten der nationalen Regierungspartner. Die Zusammenarbeit mit Regierungspartnern auf der dezentralen Ebene wurde, wo sie vorkam, allerdings besser bewertet. Ebenfalls positiv beurteilt wurden die Zusammenarbeit mit Nichtregierungspartnern, ihre Motivation und ihre Dynamik, wenngleich diese Punkte ebenfalls nur in einigen Fällen Erwähnung fanden. In einigen Fällen wurden nicht alle relevanten Akteure ausreichend und genügend partizipativ in die Planung einbezogen und/oder deren Motivation und Absichten zu wenig abgeklärt.

Über alle Evaluierungstypen hinweg tendenziell positiv beurteilt wurde die Relevanz der strategischen Ausrichtung der Vorhaben. Den Projekten wurde attestiert, sich mit wichtigen Entwicklungshemmnissen zu befassen und sowohl mit nationalen als auch internationalen Politiken im Einklang zu sein. Mit Blick auf die Programmenden wurden in den UEs aber oft eine fehlende Exit-Strategie und eine ungenügende Konsolidierung bemängelt. In Bezug auf die **Systemrelevanz** zeigt sich, dass, obwohl die Relevanz der Vorhaben gemäß DAC-Kriterium sehr positiv beurteilt wurde, die Vorhaben oft keine nachhaltige Wirkung auf der Systemebene entfalten, auch wenn es hier auch positive Beispiele gab. In einigen Fällen war zwar das individuelle Capacity Building erfolgreich, der geplante Transfer in das System scheiterte jedoch. Die Vorhaben operieren alle in einem mehr oder weniger fragilen und unberechenbaren Kontext, was eine Wirkung auf Systemebene erheblich erschwert. Absichtserklärungen seitens der Regierungspartner bezüglich Reformen finden oft keinen Niederschlag in einer entsprechenden Umsetzung. Vielfach wurde der Verlauf von Reformprozessen von Vorhaben bei der Planung und auch während der Umsetzung zu optimistisch eingeschätzt. Motivation und Bedürfnisse der verschiedenen, in die Reformprozesse involvierten Akteure wurden bei der Planung nicht ausreichend analysiert. Im Vergleich zu den UEs zeichnen die PFKs – die einen Ausschnitt aus dem aktuellen Portfolio der GIZ im Bildungsbereich wiedergeben – ein etwas positiveres Bild bezüglich der Systemrelevanz.

Im Bereich der **Geberkoordination** wurde festgestellt, dass eine Koordination mit anderen bilateralen sowie multilateralen Geldgebern in der Mehrheit der Vorhaben stattfindet. Ein zentrales Anliegen ist evaluierungstypübergreifend der Wunsch nach einer Verstärkung der Geberkoordination, um die in gewissen Fällen entstandenen Parallelstrukturen beziehungsweise Überschneidungen in Zukunft vermeiden zu können.

Monitoring und Evaluierung (im Bildungssektor)

Im Bereich der übergreifenden, strategisch besonders relevanten Fragen, die seitens der GIZ in den ToR formuliert werden, lässt sich in Bezug auf **Wirkungsnachweise** zunächst das oftmalige Fehlen der Wirkungsnachweise auf Zielgruppenebene festhalten. Zwar wird in allen Evaluierungen, welche im Rahmen der Synthese beurteilt wurden, plausibel dargelegt, welche direkten und indirekten Wirkungen ein bestimmtes Vorhaben erzielte, jedoch konnten bei weitem nicht alle Evaluierungen Wirkungen auf der Ebene der eigentlichen Zielgruppen des Vorhabens nachweisen.

Dies repräsentiert einerseits eine methodische Schwäche im Projektmanagement, da es nicht gelungen ist, im Rahmen des Monitorings Beiträge zu Verbesserungen für die Zielgruppe zu beobachten und nachzuweisen. Andererseits wurden die erwarteten Wirkungen auf die eigentliche Zielgruppe der Vorhaben in den Evaluierungen sehr oft als zu ehrgeizig beurteilt, speziell unter Berücksichtigung der Kontexte, in welchen die Vorhaben agieren.

Bezüglich der **Bildungsqualität** wird deutlich, dass diese in den evaluierten Projekten durchweg mit guten Schulleistungen in Verbindung gebracht wird. Als Messgrößen werden Schulnoten, Primarschul-Abschlussquoten und Übertrittsquoten in die Sekundarschule oder weiterführende Ausbildungsgänge herangezogen. Nicht in allen Fällen kann jedoch von guten Noten und erfolgreichen Abschlüssen darauf geschlossen werden, dass die betreffenden Schüler/-innen Kompetenzen aufbauen, die sie später nutzbringend verwenden können. In mehreren Evaluierungen werden Hinweise zur Bedeutung der Relevanz von Bildungsinhalten für die jeweiligen Schüler/-innen in ihrem Lebenskontext sowie in Bezug auf eine spätere Anschlussfähigkeit für die berufliche Aus- und Weiterbildung gemacht. Andererseits spiegelt sich darin die Reichweite der evaluierten Projekte wider, deren Ziele – gemäß der auftraggeberseitigen Vorgaben – auf der Ebene der Mittler (v.a. Bildungsministerien, Schulämter, Lehrpersonen) angesiedelt waren und Projekte nur indirekte Beiträge für Schüler leisteten.

Fachliche Aspekte im Bildungssektor

Die Bildungsprojekte, die in der Synthese untersucht wurden, können in unterschiedliche **fachliche Aspekte unterteilt werden**. Dabei werden Leistungen verschieden kombiniert. Die Erbringung der einzelnen Leistungen wird meist erfolgreich bewertet. Dabei zeigte sich, dass auch sehr vielversprechende Ansätze und Leistungen häufig erst im Zusammenspiel mit anderen Maßnahmen ihre Wirkung entfalten konnten. Besonders erfolgversprechend erwiesen sich beispielsweise die Verbindung von Lehrerfortbildungen mit der Entwicklung und Einführung neuer Lehrmittel sowie die Verwendung eines Mehrebenenansatzes. Es gilt festzuhalten, dass für die Wirkungserzielung auf Mittler- und Zielgruppenebene der kontextspezifisch adäquate Mix der einzelnen Leistungen von zentraler Bedeutung ist. Dabei erwies es sich als für das Engagement der Partner förderlich, wenn es dem Vorhaben gelang, an Bestehendes anzuknüpfen, wie beispielsweise im Tschad an die Tradition von Elternvereinen.

Im Bereich der sektorspezifischen bzw. fachlichen Erfolgs- und Misserfolgskriterien sind in allen Bildungsvorhaben die Verbesserung des Zugangs und der Bildungsqualität zentrale Anliegen. Vor allem in Subsahara-Afrika stellt der deutlich verbesserte Bildungszugang eine enorme Herausforderung für das Bildungssystem dar, da oft zahlenmäßig ausreichende und qualifizierte Lehrkräfte fehlen. Die internationale Forschung zeigt, dass eine Verbesserung des Zugangs zu Bildung Folge der Anstrengungen zur Verbesserung der Bildungsqualität relativieren und die Effekte dieser Anstrengungen herauszögern kann. Eine bessere Bildungsqualität wiederum wirkt sich auch positiv auf eine höhere Einschulungs- und geringere Abbruchrate aus. Wenn Bildung als relevant und für die Zukunft der betreffenden Schüler/-innen, ihren Familien und Gemeinden erlebt wird, steigt die Motivation zur Bildungsbeteiligung und beeinflusst den Bildungserfolg positiv. **Bildungsqualität und Zugang zu Bildung** sind entsprechend Anliegen, die eng miteinander verknüpft sind und weder parallel noch sequenziell, sondern nur systemisch bearbeitbar sind. Dies wurde in allen Bildungsvorhaben unter Berücksichtigung des Kontextes gut berücksichtigt.

In allen Grundbildungsvorhaben werden Maßnahmen zur Stärkung der **Lehreraus- und -fortbildung** umgesetzt. Es geht dabei um die Verbesserung der fachlichen und der methodisch-didaktischen Kompetenzen zur Erhöhung der pädagogischen Professionalität der Lehrkräfte. Die Entwicklung und Verteilung von neuen Lehr- und Lernmaterialien sowie von Unterrichtsmaterialien sind weitere wichtige Maßnahmen – im Idealfall wird beides kombiniert. So ist eher gewährleistet, dass Lehrkräfte diese Materialien anwenden können und dies im Unterricht auch tun. Ein gutes Beispiel dafür ist das Projekt in Ghana, bei welchem Lehrmittel in verschiedenen ghanaischen

Sprachen unter Einbezug des Personals des Bildungsdepartements sowie der Lehrerbildungsinstitutionen entwickelt wurde. Im Rahmen der Aus- und Fortbildung wird auch häufig eine Verschiebung von lehrerzentriertem zu schülerzentriertem Unterricht thematisiert (z.B. in den Projekten in Ghana, Guinea und Malawi).

Vielversprechend erweist sich die gezielte **Verwendung eines Mehrebenenansatzes**. Dieser steht dabei in engem Zusammenhang mit der oben genannten Systemrelevanz. Bei einigen UEs wurde deutlich, dass durchaus erfolgreiche Maßnahmen auf einzelnen Ebenen nur geringe oder keine breitenwirksame Wirkung entfalten konnten, weil Faktoren auf anderen Ebenen dies nicht unterstützten. Beispielsweise vermochten verbesserte Unterrichtsmaterialien und die Aus- und Weiterbildung von Lehrkräften in Pakistan keinen schülerzentrierteren Unterricht zu bewirken. Nebst einem zu geringen Vorbildungsniveau der Lehrkräfte behinderten Rahmenbedingungen wie eine schlechte Infrastruktur und zu wenig Unterrichtsmaterial in oft schlechter Druckqualität die Umsetzung eines stärker schülerzentrierten Unterrichts. Andere Beispiele in PFKs deuten darauf hin, dass Maßnahmenbündel, die idealerweise auf Mikro-, Meso- und Makroebene, zumindest aber auf zwei Ebenen ansetzen, effektiver Hebelwirkungen zeigen. Dazu zählt zum Beispiel ein Projekt in Zentralasien, in dem in den Lehrerbildungsinstitutionen der Partnerländer Innovationen mit hohem Potenzial für eine nachhaltige Verbesserung der Relevanz der Bildung und der Verwertbarkeit schulischer Lerninhalte eingeführt werden. Politischer Rückhalt für die breitenwirksame Umsetzung wird von Schlüsselakteuren auf der Partnerseite zugesichert und von qualifizierten Institutionen getragen, die durch Beratung und Kapazitätsaufbau durch das Programm in diesem Prozess unterstützt werden.

Eine Reihe evaluierter Projekte operierte in fragilen (Post-)Konfliktsituationen. Bei diesen Beispielen wurden jeweils an die Kontexte angepasste Ansätze verwendet, die der konfliktiven Situation Rechnung trugen, wie beispielsweise die Kombination von Grundbildungsförderung mit non-formaler Berufsbildung in der Demokratischen Republik Kongo oder der Förderung von „Home Schools“ für afghanische Flüchtlinge in Pakistan.

Das Thema **Geschlechtergleichberechtigung** wird in allen drei Evaluierungstypen behandelt. Für die Vorhaben südlich der Sahara, welche durch den Ex-Unternehmensteil GTZ initiiert wurden, lässt sich feststellen, dass die Thematisierung des Genderaspekts als Erfolg bezeichnet werden kann. Dies gilt besonders in Bezug auf eine erhöhte Einschulrate sowie einer geringeren Schulabbruchrate von Mädchen. Eine konstante Sensibilisierung durch die Programme hat zu diesem Erfolg beigetragen. Außerdem scheinen sich diese Erfolge auf ein System kleiner Klassen, den Einbezug insbesondere der Mütter, einen Fokus auf Mädchenklassen und gut ausgebildete Lehrerinnen zurückführen zu lassen. Als besonders wirksam erwiesen sich Ansätze, welche kontextabhängig den spezifischen Bedürfnissen von Mädchen Rechnung trugen und die Zusammenarbeit mit Ältestenräten, religiösen Führern, Elternvertretern sowie den dezentralen Bildungsstrukturen pflegten. Als gelungenes Beispiel kann dabei ein Projekt in Guinea gelten, in dem die Bildungssituation benachteiligter Bevölkerungsschichten (besonders sozial benachteiligte und lernschwache Mädchen) verbessert werden sollte; HIV & AIDS-Prävention war dabei ein zentrales Querschnittsthema. Das Projekt war dezentral aufgestellt, was sich während einer Krisenzeit, indem eine Zusammenarbeit auf nationaler Ebene nicht möglich war, als Erfolgsfaktor herausstellte. Eine weitere Wirkung im Programm „Filles éduquées réussissent“ (FIERE – gebildete Mädchen sind erfolgreich) in Guinea ist eine persönliche Einstellungs- und Verhaltensänderung der Mädchen. Diese Veränderung drückt sich bei

der Mehrzahl der befragten Mädchen darin aus, dass sie sehr viel selbstbewusster auftreten und nun lieber in die Schule gehen.

Empfehlungen

Hinsichtlich der Stärkung des **Erfolgsfaktors Kooperation** und der Stärkung der **Systemrelevanz** der Projektwirkungen geben wir folgende Empfehlungen aus:

1. Im Dialog mit dem Hauptauftraggeber, dem Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ), sollte die GIZ versuchen darauf hinzuwirken, dass realistische Ziele gesetzt und eine angemessene und finanziell abgesicherte Laufzeit geplant werden. In der Regel ist für die Umsetzung und die Konsolidierung von Reformprozessen eine **Langzeitorientierung und Konstanz** seitens des Vorhabens sinnvoll.
2. Für die Vorbereitung und Planung von Bildungsvorhaben sollte die GIZ prüfen, inwiefern entsprechende Instrumente im Auftragsmanagement – wie die Erarbeitung von **Akteurslandkarten** und die Durchführung von **Capacity Assessments** – noch konsequenter und gewinnbringender eingesetzt werden können. Zu klären wären dabei die Identifikation relevanter Stakeholder, eine Analyse von deren **Ownership und Motivation** sowie deren Rolle in bildungspolitischen Reformprozessen. Damit kann das Vorhaben mögliche Widerstände und Risiken antizipieren. Für den Bildungssektor sind die relevanten Akteure typischerweise, aber nicht ausschließlich Schüler/-innen, Eltern, Lehrpersonen, Schulverwaltungen und die jeweiligen für die Bildung zuständigen Regierungsstellen von der lokalen bis zur nationalen Ebene.
3. Die GIZ sollte bereits während der Planung von Projekten **Exit- und Konsolidierungsstrategien** in Zusammenarbeit mit den lokalen Partnern und mit der nötigen Flexibilität erstellen, um auf sich verändernde Rahmenbedingungen eingehen zu können.
4. In der Durchführung von Projekten sollten noch konsequenter als bisher **partizipative Ansätze** unter **Einbezug aller relevanten Akteure** (wie zum Beispiel die Einbeziehung von Eltern oder die Entwicklung von Unterrichtsmaterial unter Einbezug von Personal aus Bildungsverwaltungen, Lehrerbildungsinstitutionen und von Lehrern) verwendet werden, die zwar kosten- und zeitintensiv sind, aber qualitativ hochstehende und nachhaltige Resultate fördern. Gleichzeitig unterstützen solche partizipative Ansätze die gegenseitige Wertschätzung und das Vertrauen zwischen der GIZ und ihren Partnern.
5. In Fällen, in denen Reformprozesse mit staatlichen Partnern nicht vorangetrieben werden können, sollte die GIZ in ihren Projekten verstärkt mit **zivilgesellschaftlichen Nichtregierungspartnern** zusammenarbeiten. Im Bildungssektor können Nichtregierungsorganisationen (NROs) durch Mitwirkung bei der Konzeption und Finanzierung von Bildungsprojekten vitale Funktionen übernehmen. Zivilgesellschaftliche Organisationen besitzen oft eine größere Flexibilität und Veränderungsmotivation als Regierungsstellen, haben andererseits jedoch weniger direkten Einfluss auf die Systemebene. In den von uns ausgewerteten Evaluierungsberichten, in denen NROs thematisiert wurden, wurden ihre Dynamik und Motivation positiv beurteilt. Auch diese Empfehlung ist jedoch kontextabhängig. So gilt es in Gebieten mit regierungsfreundlichen sowie regierungsfeindlichen NROs im Vorfeld mögliche Nebeneffekte einer Zusammenarbeit zu bedenken.

Hinsichtlich der **Monitoring- und Evaluierungssysteme** (M+E Systeme), **insbesondere in Bildungsprojekten**, werden folgende Schritte empfohlen:

6. M+E Systeme sollten verbessert werden, um eine fortlaufende Beobachtung des Projektverlaufs und der Steuerungsentscheidungen, sowie einen **Nachweis über erzielte Wirkungen** zu ermöglichen. Für den Bildungsbereich sollte die GIZ Energie in die Erarbeitung von Lösungen zur Messung von Beiträgen auf der Zielgruppenebene investieren, auch wenn diese in der Planung der Vorhaben außerhalb des Verantwortungsbereichs der GIZ liegt. In diesem Zusammenhang sollte die Sinnhaftigkeit von Indikatoren immer wieder überprüft werden. Im GIZ-Kontext sind Indikatoren sinnvoll, die über Veränderungen auf der Ebene von Mittlern Schlüsse auf Verbesserungen der Lernleistungen bei Schülerinnen und Schülern zulassen. Auf Zielgruppenebene bieten sich besonders regionale Lernstandsmessungen wie beispielsweise vom Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) oder im Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC) an, welche den regionalen Kontexten gerecht werden.
7. Wir empfehlen, dass Innovationen auf ihre breite Anwendbarkeit und eine weitere Verbreitung im Bildungssektor anderer Länder überprüft werden, der Transfer von guten Beispielen mit Anpassungen für den anderen Kontext kann signifikante Synergien bewirken. Die bestehenden **M+E Systeme** sind nicht ausreichend geeignet dafür, die **Kontextfaktoren**, unter denen bestimmte methodische Ansätze ihre Wirkungen entfalten können, empirisch zu erfassen. Dies wäre aber nötig, um die Übertragbarkeit von Innovationen realistisch einschätzen zu können. Dafür wird eine wissenschaftliche Begleitforschung empfohlen. Eine Verbindung des M+E Systems mit der Akteurslandkarte wäre zu prüfen.

Bezüglich der fachlichen Aspekte wird der GIZ Folgendes empfohlen:

8. Sofern **Bildungsqualität** mit **Lernleistungen** und diese wiederum mit der Kompetenz der Schüler gleichgesetzt werden (wie in zahlreichen im Bildungswesen evaluierten Projekten der Fall), empfehlen wir, dass die GIZ thematisiert, woran Lernleistungen sinnvoll gemessen werden können. Vorsicht ist geboten, wenn dazu ausschließlich Schulnoten herangezogen werden. Die Aussagekraft von Schulnoten ist in der Bildungsforschung immer wieder Gegenstand von Debatten, darüber hinaus spielt es eine wichtige Rolle, wie Schulnoten zustande kommen. Ob Schüler/-innen in der Lage sind ihr Wissen nutzbringend anzuwenden und ob ihre Kompetenzen den Anschluss an die berufliche Aus- und Weiterbildung ermöglichen, ist somit nicht automatisch anhand von Schulnoten abzulesen. Kompetenzbasierte Lehrpläne, Standards sowie Prüfungen können ein gangbarer Weg der Verbesserung sein; ein Beispiel für eine solche Vorgehensweise stellen die Bildungsstandards in Deutschland dar (vgl. z.B. Klieme, 2007; Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2004).
9. **Zugang zu Bildung und Bildungsqualität** stehen meist in einem engen Zusammenhang. Wir empfehlen beide Anliegen in Bildungsvorhaben weiterhin konsequent im System und nicht sequenziell im Blick zu haben.
10. Sofern der jeweils lokale Kontext einen entsprechenden Rückhalt dafür bietet, empfehlen wir weiterhin in der Beratung von Partnern im Bildungswesen auf eine Veränderung **von stark lehrerzentrierten hin zu mehr schülerzentrierten Unterrichtsformen** hinzuwirken,

insbesondere in Kombination und Ergänzung mit (direktiven) Orientierungen und Ordnungsstrukturen (vgl. Hattie, 2009). Damit dieses Potenzial wirksam wird, sind insbesondere folgende Faktoren bei Lehrerfortbildung und der Entwicklung von Lehrmaterialien zu berücksichtigen:

- Ein Wechsel von lehrerzentrierten hin zu schülerzentrierten Unterrichtsformen bedingt tiefgreifende Veränderungen der Haltung und des Selbstverständnisses von Lehrern, was wiederum eng mit Ownership-Fragen verknüpft ist.
- Qualitativ guter schülerzentrierter Unterricht benötigt noch besser ausgebildete Lehrkräfte als ein lehrerzentriertes Setting.
- Erfolgreiche schülerzentrierte Unterrichtsformen benötigen entsprechend angepasste Lehrpläne, Unterrichtsmaterialien und Prüfungssysteme.

11. **Unterrichtsmaterial** sollte weiterhin nach Möglichkeit partizipativ und in hoher Qualität entwickelt werden, mit Erprobungsschlaufen in der Praxis (Mikroebene), gekoppelt mit Lehrerbildung und Organisationsentwicklung / Lehrplanentwicklung auf Makroebene.

1 Einleitung

1.1 Übersicht

Das Institut für internationale Zusammenarbeit in Bildungsfragen (IZB) und das Institut für Bildungsmanagement und Bildungsökonomie (IBB) der Pädagogischen Hochschule Zug verfassten im Auftrag der Deutschen Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH eine Querschnittsanalyse (QSA) bestehend aus einer Meta-Evaluierung und einer Synthese von Evaluierungen von Projekten der GIZ im Sektor Bildung. Gemäß den Terms of Reference (ToR) verstehen wir darunter eine Untersuchung der methodischen Qualität von Evaluierungen (Meta-Evaluierung) sowie eine Zusammenfassung der wichtigsten Erkenntnisse aus den Einzelevaluierungen (Synthese).

Wie in den ToR dargestellt, sehen wir die Meta-Evaluierung als Grundlage „der Einschätzung der Belastbarkeit von Aussagen aus den uns zur Verfügung gestellten Einzelevaluierungen, was wiederum eine Voraussetzung für nachvollziehbare und möglichst robuste Aussagen im Rahmen der Synthese darstellt“ (ToR, S. 2).

Datengrundlage bilden zehn unabhängige Evaluierungen (UEs²), zwölf Projektfortschrittskontrollen (PFKs) sowie zwei dezentrale Evaluierungen (DEs), welche die Stabsstelle Monitoring und Evaluation hat durchführen lassen. Eine ausführliche Übersicht der Datengrundlage befindet sich im Anhang (Kapitel 6.2 auf Seite 66).³

In der vorliegenden QSA stellen wir die Datengrundlage und Methoden zur Durchführung der Meta-Evaluierung kurz vor; dabei gehen wir auch auf die Grenzen der Untersuchung ein. Anschließend werden die Ergebnisse der Meta-Evaluierung zusammengefasst präsentiert, detaillierte Datenblätter aller untersuchten Dokumente befinden sich im Anhang.

Im darauf folgenden zweiten Teil (Synthese) erläutern wir zunächst kurz die Zielsetzung und stellen die Methoden zur Erstellung der Synthese vor. Anschließend werden die Ergebnisse, sortiert nach Themen und nach Evaluierungstyp (UE, PFK, DE) präsentiert. Dies ermöglicht ein fokussiertes Lesen, beispielsweise ausschließlich zu Ergebnissen, welche PFKs betreffen oder ausschließlich zu Ergebnissen aus dem Bereich der manageriale Erfolgs- und Misserfolgskriterien. Im Anschluss ziehen wir ein Fazit, basierend auf den zuvor dargestellten Ergebnissen. Das Fazit orientiert sich an zuvor erarbeiteten Themen (z.B. „Manageriale Erfolgs- und Misserfolgskriterien“) und ist größtenteils evaluationstyp-übergreifend. Dort, wo zwischen den Evaluationstypen Unterschiede deutlich werden, wird im Text differenziert auf unterschiedliche Typen eingegangen. Ein Kapitel mit Empfehlungen schließt die QSA ab.

²In diesem Dokument lautet der Plural von UE UEs, der Plural von PFK PFKs und der Plural von DE DEs.

³ Die UEs und die PFKs beziehen sich auf unterschiedliche Zeiträume. Die Schluss- und Ex-Post Evaluierungen beziehen sich auf Vorhaben, die zwischen Jahr 2012 und Jahr 2013 durchgeführt wurden. PFKs sind Zwischenevaluierungen laufender Maßnahmen, entsprechend spiegeln sie ein aktuelleres Portfolio wider. In der Synthese werden UEs und PFKs separat behandelt.

1.2 Datengrundlage

Datengrundlage bilden zehn unabhängige Evaluierungen (UEs), zwölf Projektfortschrittskontrollen (PFKs) sowie zwei dezentrale Evaluierungen (DEs). Eine ausführliche Übersicht der Datengrundlage befindet sich im Anhang (S.66).

Die der Meta-Evaluierung und Synthese zugrunde liegenden PFKs, UEs und DEs wurden gemäß den Kriterien *Zuordnung zum Sektor Bildung, Region, Projekttyp, Zeitraum* (vor/nach der GIZ-Fusion) sowie der *Durchführungsorganisation* gegliedert. Die Gliederung zur Zuordnung zum Sektor Bildung wurde wiederum in vier Unterkategorien aufgeteilt. Der Unterkategorie *Gesamter Sektor Bildung* wurde ein Vorhaben zugeordnet, sofern es alle Bildungssektoren behandelt. Dies sind Berufsbildung, Hochschulbildung, frühkindliche Bildung sowie Grundbildung und Sekundarbildung. Der Unterkategorie *Teilaspekt des Sektors* wurde ein Vorhaben zugeordnet, wenn das Hauptaugenmerk auf einem oder mehreren Bildungssubsektoren liegt. So wurde das Vorhaben in Ghana mit seinem Hauptaugenmerk auf Grundbildung beispielsweise der Kategorie *Teilaspekt des Sektors* zugeteilt. Der Unterkategorie *Subkomponente* wurde ein Vorhaben zugeschrieben, sofern weder der gesamte Bildungssektor behandelt wird noch ein Teilaspekt des Sektors im Zentrum der Evaluierung steht. Der letzten Unterkategorie, *Wirkung auf den Sektor Bildung*, wurde ein Projekt schließlich zugeteilt, wenn die vorangehenden Fragen alle mit Nein beantwortet wurden, das Vorhaben demzufolge kein Bildungsprojekt ist, es jedoch Wirkungen auf den Sektor generieren kann. Nachfolgend werden die Ergebnisse der genaueren Bestimmung der Stichprobe kurz zusammengefasst. Die ausführliche Zuordnung der einzelnen Vorhaben zu den jeweiligen Kategorien befinden sich in tabellarischer Form im Anhang auf Seite 78.

Im Hinblick auf die Aspekte der Vergleichbarkeit lassen sich bei den UEs folgende Resultate extrahieren: Kein unabhängig evaluiertes Vorhaben nimmt den gesamten Sektor Bildung in den Blick. Bei acht Vorhaben handelt es sich schwerpunkttechnisch um Grundbildungsprojekte (GHA, GUI, HGP, INA, NAM, PER, CHA), im Fall von Pakistan um ein Grundbildungsvorhaben mit zusätzlichen Elementen der Berufsbildung. Nur im Vorhaben in Zentralamerika handelt es sich im engeren Sinne nicht um ein Bildungsvorhaben, sondern um ein Projekt, welches das öffentliche Budgetmanagement im Bildungssektor im Visier hat, jedoch als Subkomponente Grundbildungsinhalte aufweist.

Die regionale Streuung der Vorhaben ist relativ groß. In der Region Asien/Pazifik (INA) sowie im Nahen Osten (PAK) befindet sich in der Stichprobe nur je ein Vorhaben. In Lateinamerika sind es vier Vorhaben (HGP, PER, ZA, ZAM), südlich der Sahara vier (GHA, GUI, NAM, CHA). Bezüglich des Projekttypen handelt es sich in sieben Fällen um bilaterale und in zwei Fällen um regionale Vorhaben (HGP, ZA). Der Durchführungszeitraum der Vorhaben liegt größtenteils vor der Fusion der drei ehemaligen Unternehmensteile zur heutigen GIZ im Jahr 2011. Sieben Vorhaben wurden vor der Fusion durchgeführt und auch abgeschlossen. Die Vorhaben in Honduras, Guatemala und Peru sowie in Zentralamerika wurden erst nach der Fusion abgeschlossen.

Bei den Durchführungsorganisationen handelt es sich bei sieben Vorhaben (GHA, GUI, INA, NAM, PAK, CHA, PER), welche vor der Fusion abgeschlossen wurden, um GTZ Projekte. Bei den beiden Vorhaben, deren Projektdauer die Fusion überdauerte, handelt es sich ebenfalls um InWEnt Projekte (HGP, ZA). Des Weiteren gilt es festzuhalten, dass es sich bei allen Vorhaben südlich der

Sahara um Vorhaben des Ex-Unternehmensteils GTZ handelt und die Vorhaben in Lateinamerika mit Ausnahme des Vorhabens in Peru ausschließlich von InWEnt durchgeführt wurden.

Bei den zwölf Projektfortschrittskontrollen nimmt nur das Sektorvorhaben den gesamten Sektor Bildung ins Visier. Bei den weiteren elf PFKs handelt es sich bei fünf Vorhaben hauptsächlich um Grundbildungsprojekte (AFG, GUI, HON, MAW, SRI). Im Kongo sowie in Mosambik findet eine Kombination aus Grundbildungs- und Berufsbildungsmaßnahmen statt. Drei Vorhaben setzen zudem nicht direkt an einem bestimmten Teilsektor an (KOS, MAW, ZAs), haben jedoch einen bestimmten Bildungssektor als Subkomponente. Im Kosovo und in Sri Lanka handelt es sich bei der Subkomponente um Grundbildung, in Zentralasien um eine Kombination aus Grund- und Berufsbildung. Die beiden einzigen Vorhaben im Rahmen der PFKs, welche keinen Teilaspekt des Sektors Bildung aufweisen, auch nicht als Subkomponente, sind die Vorhaben in Benin und Mauretanien. Das Vorhaben in Mauretanien zielt in erster Linie auf gute Regierungsführung ab; durch Fortbildungsmaßnahmen zu mehr Transparenz, Dezentralisierung und Staatshaushalt profitiert nicht nur die Regierungsführung und Dezentralisierung, sondern kann auch der Sektor Bildung durch eine verbesserte Planung und Finanzierung positiv beeinflusst werden. Im Benin handelt es sich um ein Kapazitätsentwicklungsvorhaben für Schulämter. Hinsichtlich der räumlichen Streuung der Vorhaben zeigt sich bei den PFKs ein ähnliches Bild wie bei den UEs. Sechs Projekte sind südlich der Sahara verortet (BEN, GUI, COD, MAW, MTN, MOZ). Die weiteren Vorhaben verteilen sich je einmal auf die Regionen Asien/Pazifik (SRI), Lateinamerika (HON), Naher Osten (AFG), Südosteuropa (KOS) und Zentralasien (ZAs). Beim Projekttyp handelt es sich in zehn Fällen um bilaterale Vorhaben, um ein Sektorvorhaben sowie um ein regionales Vorhaben (ZAs). Das Vorhaben SV wurde als einziges erst nach der Fusion begonnen und ist somit ein reines GIZ Vorhaben. Der Durchführungszeitraum der evaluierten Projektphasen im Rahmen der PFKs beginnt und endet nur im Fall des GTZ-Vorhabens in Guinea vor der Fusion der drei ehemaligen Unternehmensteile der GIZ. Bei den restlichen Vorhaben handelt es sich um Projekte, welche fusionsübergreifend liefen bzw. immer noch laufen. Von diesen zehn fusionsübergreifenden Vorhaben wurden acht durch die GTZ initiiert (AFG, BEN, COD, MTN, MOZ, SRI, ZAs, HON) und vier durch InWEnt, (KOS, MAW, ZA, HGP).

Bei den beiden dezentral evaluierten Vorhaben handelt es sich schließlich nicht um klassische Bildungsprojekte, wie sie im Rahmen von UEs und PFKs auftauchen. Die beiden Vorhaben haben weder den gesamten Sektor Bildung im Blick, noch steht ein einzelner Teilaspekt des Bildungssektors im Projektfokus. In Zentralamerika ist der Bildungsaspekt in Gestalt von Grundbildung auf der Subkomponentenebene zu lokalisieren. Im Kosovo handelt es sich um ein Jugendförderungsvorhaben, welches jedoch auf den Sektor Bildung wirkt, da Fragen der Grund- und Berufsbildung mit den Jugendlichen thematisiert werden. Eine Visualisierung der bereinigten Stichprobe befindet sich im Anhang auf Seite 78.

2 Metaevaluierung

2.1 Methodik der Metaevaluierung

Um fundierte Aussagen über die Qualität von Evaluierungen machen zu können, müssen die Bewertungsergebnisse und Folgerungen nachvollziehbar auf empirisch gewonnenen qualitativen und/oder quantitativen Daten beruhen (vgl. Beywl, 2008, S. 15). Gleiches gilt für das Identifizieren von Erfolgs- und Misserfolgskriterien im Rahmen der durchzuführenden Synthese. Es wurde daher zunächst eine Meta-Evaluierung durchgeführt, anschließend erfolgte eine Synthese all jener Evaluierungen, die im Rahmen der Meta-Evaluierungen als solide erachtet wurden.

Die in den ToR genannten Evaluierungsfragen wurden entlang verschiedener Themenschwerpunkte gebündelt. Außerdem ergänzten wir die Evaluierungsfragen basierend auf unseren Erfahrungen aus der Beratungstätigkeit in der internationalen Zusammenarbeit und basierend auf Gesprächen mit der GIZ um einen weiteren Themenschwerpunkt (Vorhandensein von Empfehlungen in den Evaluierungen, Qualität dieser Empfehlungen).

Die Bearbeitung der Evaluierungsfragen erfolgte kriteriengeleitet, das heißt, dass ein Kriteriensystem entwickelt und anschließend auf die Evaluierungen angewandt wurde. Inhaltliche Basis für dieses Kriteriensystem stellen die ToR der GIZ für die QSA des Sektors Bildung dar; sie enthalten eine Reihe von Fragen (siehe Kapitel 6.3 auf S. 68 im Anhang), aus welchen Kriterien abgeleitet wurden. Die methodische Basis stellt ein Analyseraster für die Fragen bezüglich der Meta-Evaluierung dar (siehe Kapitel 6.4 auf S. 70 im Anhang). In diesem Raster wurden für die in den ToR genannten Fragestellungen Indikatoren definiert, d.h. die Fragen wurden in beurteilbare Einzelaspekte zerlegt. So wird beispielsweise die Frage, ob die methodische Vorgehensweise transparent dargestellt wurde, in drei überprüfbare Gesichtspunkte unterteilt: Wurden die Arbeitsschritte beschrieben? Sind Daten und Angaben vollständig vorhanden? Wurde die Methodenwahl begründet?

Methodisch stellt das Ableiten von Indikatoren und deren Verwendung in einem Analyseraster oder einer Checkliste eine konventionelle und bewährte Vorgehensweise bei Evaluierungen dar (vgl. beispielsweise Scriven, 1991, S. 204ff); auch bei anderen Meta-Evaluierungen innerhalb der GIZ kamen Varianten dieses Verfahrens zum Einsatz (beispielsweise in der von Rambøll Management Consulting durchgeführten GIZ-Meta-Evaluierung Gesundheit, vgl. Rambøll 2013). Ein signifikanter Spielraum bleibt hingegen bei der Gewichtung einzelner Evaluierungsfragen und Indikatoren für die Gesamtbewertung einer Evaluierung. Die Gewichtung der einzelnen Kriterien ist gleichwertig und wurde im Dialog mit der GIZ abgesprochen. Die Analyseebenen der Meta-Evaluierung lauten im Detail: Berücksichtigung der Wirkungsebenen (1.), Transparenz im methodischen Vorgehen (2.1), Angemessenheit von Design und Datenerhebung (2.2), Adäquatheit der Datenbasis (2.3), Verwendung der OECD-DAC Kriterien (3.) sowie Empfehlungen (4).

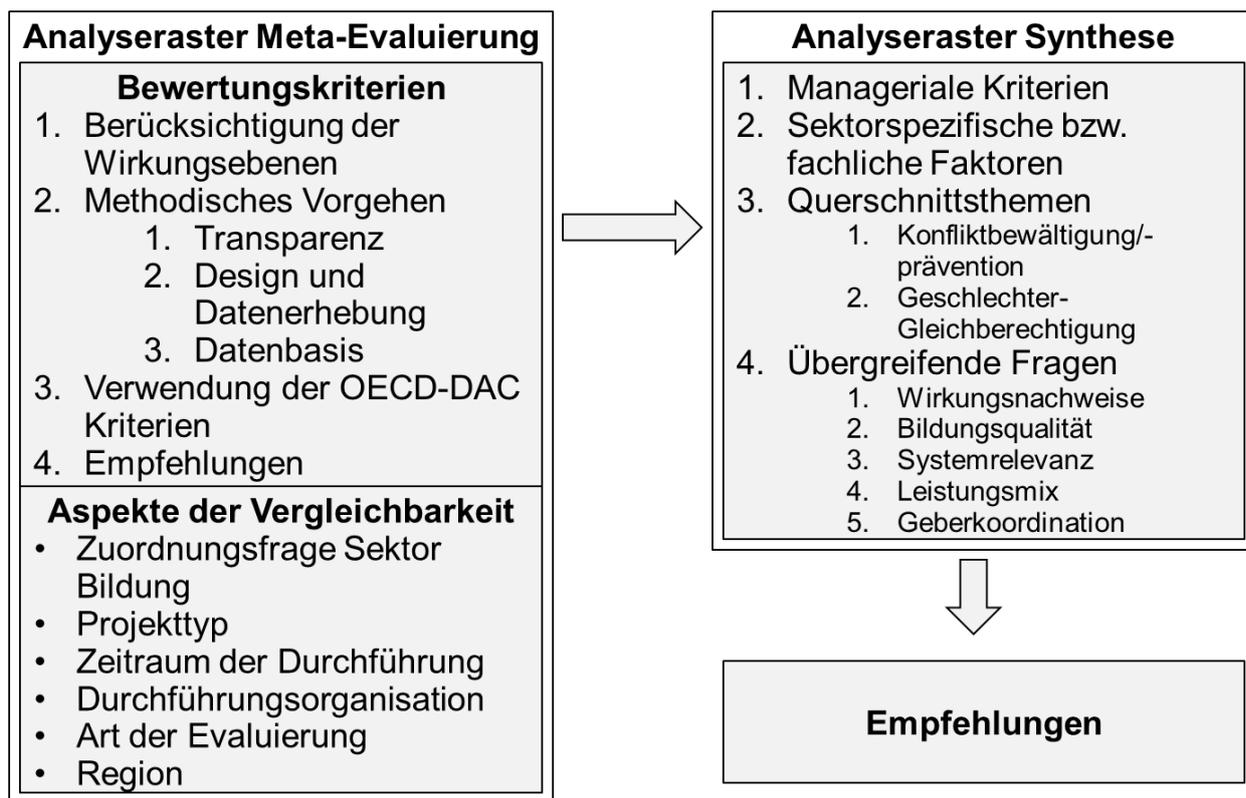


Abbildung 1: Methodisches Modell QSA (Analyseraster Meta-Evaluierung links)

Für die durchzuführende Meta-Evaluierung wurde ein Bewertungssystem entwickelt, welches zunächst Fragen nach der Robustheit der einzelnen Evaluierungen beantwortet, ergänzt mit Ergebnissen zu Fragen der Zuordnung der Evaluierungen, welche eher informativen Charakter haben. Die Robustheit aller Evaluierungen prüften wir entlang der zuvor erwähnten Kriterien, bezeichnet als Bewertungskriterien (in der Abbildung 1 links, unter der Überschrift „Bewertungskriterien“). Diese Kriterien sind jeweils in Indikatoren unterteilt (in der Abbildung 1 der Übersicht wegen nicht aufgeführt, jedoch im Anhang auf S. 70 zu finden). Für jedes Bewertungskriterium vergaben wir ein Rating basierend auf Punkte-Skalen. Basierend auf den Bewertungen der Bewertungskriterien vergaben wir darüber hinaus eine Gesamtbewertung. Zur Visualisierung der Bewertungen wurde ein Rating-basiertes Ampelsystem eingesetzt. In dem Ampelsystem sind einerseits jeweils alle Evaluierungen eines Typs aufgeführt (horizontale Ebene), andererseits alle Bewertungskriterien (vertikale Ebene). Dies ermöglicht es, einen schnellen Überblick über die Ergebnisse der Bewertungen zu erhalten. So werden Stärken und Schwächen der Evaluierungen wie auch über Stärken und Schwächen der Erfüllung einzelner Kriterien über Evaluierungen hinweg rasch ersichtlich. Eine ausführliche Darstellung des Rating-Systems sowie des damit verbundenen Ampel-Systems ist im Anhang im Kapitel 6.1 auf S. 61 zu finden.

Bezüglich der Verwendung der OECD-DAC Kriterien in den einzelnen Evaluierungen beurteilten wir außerdem, ob diese aus unserer Sicht angemessen, zu streng oder zu mild waren. Das Urteil der Angemessenheit der Benotungen basiert auf der Kohärenz der gemachten Aussagen in den

Evaluierungen. Jeweils zwei Personen lasen unabhängig voneinander die Ausführungen zu den jeweiligen DAC-Kriterien sowie die zugehörigen Benotungen und bewerteten diese als zu mild, zu streng oder angemessen. In Fällen, in denen die beiden Personen nicht zum gleichen Urteil kamen, wurde eine dritte Person hinzugezogen. Wurde eine Benotung als zu mild bewertet, sind wir der Ansicht, dass die Summe der gemachten Aussagen der Evaluatoren über die tatsächlich erzielte Wirkung in einem Missverhältnis zur Benotung anhand der GIZ-Richtlinien steht. Da die OECD-DAC Kriterien in der Synthese dezidiert behandelt werden, finden sich Aussagen über unsere Beurteilungen hierzu ebenfalls im Synthese-Teil. Die Datenblätter zu den Evaluierungen im Anhang enthalten all unsere Einschätzungen zur Angemessenheit der Bewertung der DAC-Kriterien.

Eine methodologische Begrenzung der Meta-Evaluierung ergibt sich allerdings durch die Datengrundlage selbst: Einzig die finalen Evaluierungsdokumente werden analysiert, was bedeutet, dass beispielsweise Inception Reports der UEs oder bisherige Kommentierungen nicht berücksichtigt wurden.

Darüber hinaus enthält das Raster Fragen zu Aspekten der Zuordnung und Vergleichbarkeit (in der Abbildung 1 zusammengefasst unter „Aspekte der Vergleichbarkeit“, eine ausführliche Auflistung befindet sich im Anhang auf S. 70). Diese Fragen betreffen nicht die Robustheit der einzelnen Evaluierungen, liefern jedoch wichtige Informationen im Hinblick auf die Verortung der einzelnen Evaluierungen und geben auch Hinweise auf die Vergleichbarkeit der Evaluierungen im Hinblick auf die Synthese.

Während das Analyseraster Meta-Evaluierung für jede Evaluierung erstellt wurde (in der Abbildung 1 links), wurde das Analyseraster Synthese nur bei den Evaluierungen angewandt, die als robust bewertet wurden. Das Analyseraster Synthese fasst die aggregierten Daten zu den einzelnen Kriterien in Bezug auf die drei Evaluierungstypen (UE, PFK und DE) zusammen (in der Abbildung 1 rechts). Mehr Informationen hierzu finden sich im zweiten Teil dieses Berichts, der Synthese, ab Seite 23.

Die Aussagekraft und Belastbarkeit unserer Analysen im Rahmen der Synthese wird beeinflusst durch die Heterogenität der Datengrundlage. Die verschiedenen Evaluierungen unterscheiden sich neben den Orten, in denen sie entstanden sind, u.a. bezüglich Inhalt, Umfang, Art der Evaluierung, Sprache und Vorgehensweise.

2.2 Ergebnisse

	UE (10)										Durchschnitt pro Kriterium
	GHA	GUI	HGP	INA	NAM	PAK	PER	CHA	ZA	SV	
Wirkungskette (max. 15 Punkte)	12	11	9	11	15	12	13	10	14	15	12
Methodische Transparenz (max. 5 Punkte)	3	3	4	5	5	4	5	5	3	5	4
Design und Datenerhebung (max. 3 Punkte)	3	3	1	2	3	2	2	2	2	2	2
Datenbasis (max. 5 Punkte)	4	4	3	4	4	4	5	2	4	5	4
DAC-Kriterien (max. 15 Punkte)	11	12	13	12	10	14	12	12	13	13	12
Empfehlungen (max. 8 Punkte)	7	8	8	8	8	8	6	7	7	8	8
Evaluierungsbewertung (max. 18 Punkte, min. 6 Punkte)											
	15	15	15	15	17	16	15	15	17	17	

Tabelle 1: Zusammenfassung der Ergebnisse der Meta-Evaluierung der UEs

Wie die Tabelle deutlich macht, fallen die Ergebnisse der Meta-Evaluierung der UEs positiv aus: In allen Bewertungskriterien werden mindestens ausreichende Ergebnisse erzielt. Hervorzuheben sind die methodische Transparenz, die verwendeten Datenbasen sowie die Empfehlungen; in diesen Kriterien erreichen die untersuchten UEs beinahe ausnahmslos sehr gute Ergebnisse. Zu erwähnen sind ferner die besonders guten Ergebnisse der UEs für Namibia sowie Zentralamerika. Ebenfalls sehr gute Ergebnisse erreicht die UE SV; als UE eines Sektorvorhabens ist die Vergleichbarkeit mit UEs „klassischer“ Vorhaben jedoch eingeschränkt. Etwas Potenzial zur Verbesserung besteht bezüglich der Wirkungsketten, der Angemessenheit der Designs und Datenerhebungen sowie der Verwendung der OECD-DAC Kriterien gemäß den Vorgaben der GIZ.

PFK (12)													Durchschnitt pro Kriterium
AFG	BEN	GUI	HON	COD	KOS	MAW	MTN	MOZ	SRI	SV	ZAs		
Wirkungskette (max. 13 Punkte)	6	6	7	4	7	0.	0.	4	10	0.	6	6	5
Methodische Transparenz (max. 5 Punkte)	0	2	1	3	3	0.	1	3	3	1	4	1	2
Design und Datenerhebung (max. 3 Punkte)	3	3	3	3	3	0.	3	3	3	2	3	3	3
Datenbasis (max. 5 Punkte)	2	4	3	3	3	0.	3	3	4	4	3	2	3
DAC-Kriterien (max. 15 Punkte)	0.	12	12	10	13	0.	14	9	14	6	14	11	10
Empfehlungen (max. 8 Punkte)	0	3	5	6	7	4	6	3	5	7	7	7	5
Evaluierungsbewertung (max. 18 Punkte, min. 6 Punkte)													
	9	13	12	12	15	7	12	12	15	10	15	12	

Tabelle 2: Zusammenfassung der Ergebnisse der Meta-Evaluierung der PFKs

Bezüglich der PFKs ist zunächst festzuhalten, dass die Formulierung einer Wirkungshypothese zum Zeitpunkt der Erstellung der hier untersuchten PFKs keine Anforderung darstellte. Bei der Berechnung der Beurteilungen im Bereich der Wirkungsketten haben wir daher auf eine Bewertung der Wirkungshypothese verzichtet, was dazu führt, dass im Bereich der Wirkungsketten bei PFKs maximal 13 statt wie bei UEs und DEs 15 Punkte erreichbar sind. Es zeigt sich hier ein gemischteres Bild: Drei der untersuchten PFKs erreichen nicht die Voraussetzungen, um als robust eingestuft zu werden und somit im Rahmen der Synthese weiter untersucht zu werden. Bei diesen drei PFKs handelt es sich um die Vorhaben in Afghanistan, Kosovo und in Sri Lanka. PFK-übergreifend lässt sich festhalten, dass im Bereich der Wirkungsebenen sowie im Bereich der methodischen Transparenz Spielräume der Verbesserung existieren, während Design und Datenerhebung in den meisten Fällen von uns als angemessen erachtet wurden.

DE (2)	KOS ZAM		Durchschnitt pro Kriterium
	Wirkungskette (max. 15 Punkte)	8	
Methodische Transparenz (max. 5 Punkte)	4	5	5
Design und Datenerhebung (max. 3 Punkte)	2	2	2
Datenbasis (max. 5 Punkte)	5	3	4
DAC-Kriterien (max. 15 Punkte)	5	5	5
Empfehlungen (max. 8 Punkte)	6	8	7
Evaluierungsbewertung (max. 18 Punkte, min 6 Punkte)			
	13	12	

Tabelle 3: Zusammenfassung der Ergebnisse der Meta-Evaluierung der DEs

Bezüglich der DEs lassen sich aufgrund der geringen Stichprobengröße (N=2) keine Trends ausmachen. Die beiden untersuchten DEs weisen Verbesserungspotenziale im Bereich der DAC-Kriterien auf, im Bereich der methodischen Transparenz werden besonders gute Bewertungen erzielt.

2.3 Aspekte der Auswertung und Vergleichbarkeit

Der zweite Teil der Meta-Evaluierung soll die einzelnen Evaluierungen auf ihre Vergleichbarkeit überprüfen, wie es in der Abbildung 1 ersichtlich wird. Dies ist von zentraler Bedeutung für die anschließend stattfindende Synthese.

Wie im Kapitel zur Stichprobenbeschreibung bereits aufgezeigt wurde, ist eine größere Gruppierung von einzelnen Vorhaben nicht möglich. Der Hauptgrund hierfür liegt in der geringen Evaluierungsanzahl pro Evaluierungstyp und Projektregion sowie an den sehr unterschiedlichen Problemstellungen in den verschiedenen Projektländern (siehe hierzu auch die Tabelle im Anhang auf Seite 78). Im Rahmen der Synthese kommen zusätzliche Herausforderungen bezüglich der Vergleichbarkeit der einzelnen Evaluierungen hinzu. Selbst bei Evaluierungen von Projekten aus denselben Regionen und/oder mit derselben Durchführungsorganisation kommen eine Vielzahl an weiteren Faktoren hinzu wie z.B. unterschiedliche politische Kontexte, verschiedene Mittlergruppen, die Kombination verschiedener Partner und verschiedene Innovationsansätze, um nur einige zu nennen. All diese Faktoren variieren so stark, dass die einzelnen Vorhaben nur sehr beschränkt vergleichbar sind. Auf Grund dieser Gegebenheiten, welche sowohl aus der Meta-Evaluierung als auch aus der Synthese hervorgehen, ist ein ursprüngliches Teilziel der Synthese, pro Evaluierungstyp verallgemeinerbare Aussagen zu generieren, nicht erfüllbar. Gleichwohl werden in der

folgenden Synthese auf Basis systematischer Vergleiche übergreifende Aspekte sowie inhaltliche Tendenzen herausgearbeitet.

Die relativen Beiträge verwendeter Instrumente (also Entsendung von Langzeit- und Kurzzeitpersonal, sog. Human Capacity Development (HCD)-Fortbildungsmaßnahmen, Entsendung von Entwicklungshelfern, und andere mehr) zum Erfolg bzw. Misserfolg von Vorhaben lassen sich basierend auf der vorhandenen Datengrundlage nicht seriös beurteilen. Als Folge werden in der Synthese keine Aussagen zum Instrumentenmix getroffen.

3 Synthese

Zielsetzung der Synthese ist, evaluierungsübergreifende Erkenntnisse möglichst aussagekräftig und anschaulich zu aggregieren, um schlussendlich zu praxisnahen Empfehlungen zu gelangen.

3.1 Methodik der Synthese

Für die Synthese wurden nur die Evaluierungen berücksichtigt, die bei der Prüfung der Robustheit ein ausreichendes Ergebnis vorwiesen.⁴ Mit diesem Vorgehen wurde sichergestellt, dass methodisch unzureichende Evaluierungen, deren Ergebnisse im bedeutsamen Maße fragwürdig wären, nicht in die Synthese einfließen. Als methodisches Grundgerüst für die Erstellung der Synthese wurde erneut ein detailliertes Analyseraster verwendet, welches Indikatoren und Zusammenstellungen enthält, die aus den Fragen und Anliegen der Synthese in den ToR abgeleitet wurden (siehe S. 68 im Anhang). Aus allen untersuchten Evaluierungen wurden dazu zentrale Textelemente rund um die Erfolgsaspekte von Capacity Works extrahiert. Die Textelemente wurden anschließend induktiv codiert, mit dem Ziel, übergeordnete Themen zu identifizieren. Die codierten Textpassagen wurden anschließend geclustert und unter den Bereichen „Manageriale Erfolgs- und Misserfolgskriterien“, „sektorspezifische bzw. fachliche Erfolgs- und Misserfolgskriterien“ zusammengefasst. Diese stellen somit auch Kapitel im Synthese-Teil dieser QSA dar. Ferner ergaben sich aus den ToR einige übergreifende Themen wie beispielsweise die Systemrelevanz oder die Bildungsqualität. Darüber hinaus wurden Querschnittsthemen wie z.B. Konfliktbearbeitung/Konfliktprävention seitens der GIZ festgelegt. Für die Synthese wurden die Evaluierungen daher im Hinblick auf Passagen mit Relevanz für die übergreifenden Fragen sowie für die Querschnittsthemen durchsucht (deduktives Codieren). So entstanden Kapitel zu den Querschnittsthemen wie auch zu den übergreifenden Fragen. Die Daten wurden für die drei Evaluierungstypen (UE, PFK, DE) jeweils separat ausgewertet.

⁴ Die notwendige Robustheit nicht erreicht haben, wie wir im Kapitel 1.3 gesehen haben, die drei PFKs Afghanistan, Kosovo sowie Sri Lanka.

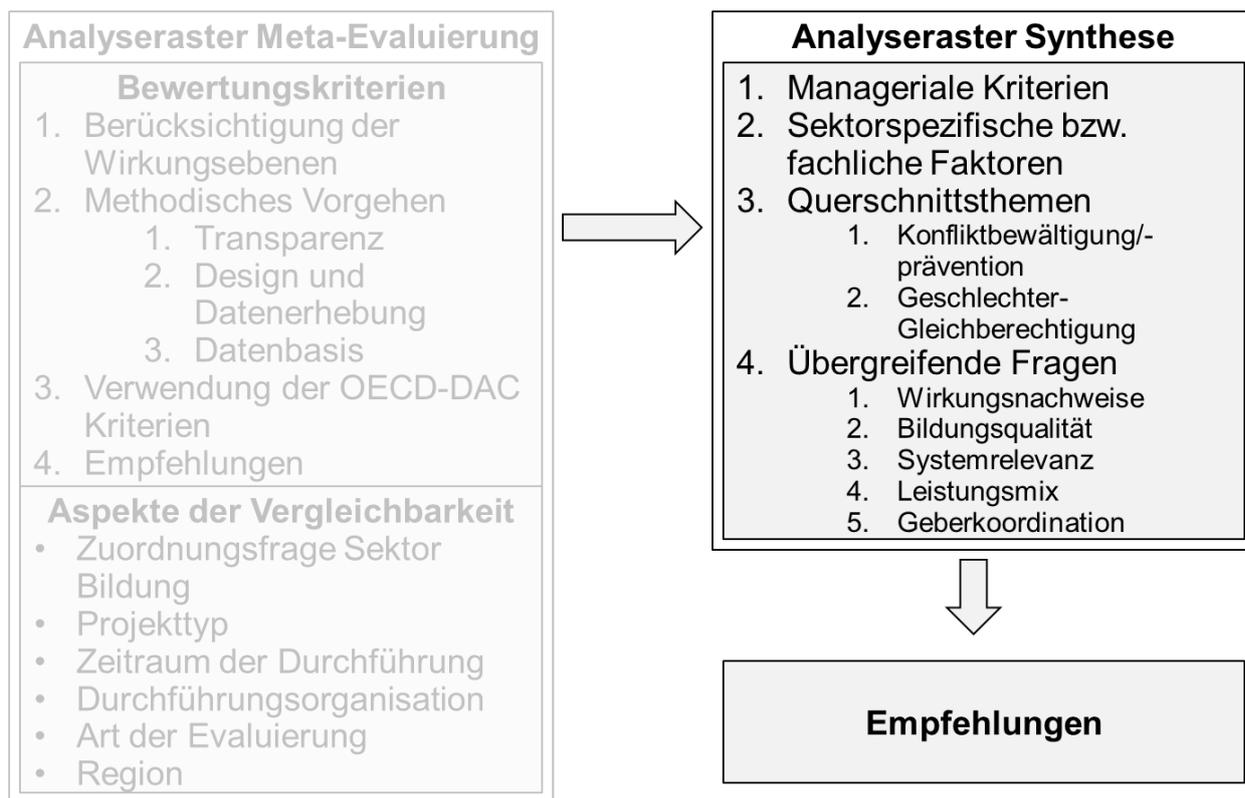


Abbildung 2: Methodisches Modell QSA (Analyseraster Synthese rechts)

Das Analyseraster Synthese fasst die aggregierten Daten entlang mehrerer Themenkomplexe sowie entlang der drei Evaluierungstypen zusammen (UE, PFK und DE, siehe hierzu Abbildung 2).

Die OECD-DAC Kriterien werden im Rahmen der Synthese zwar nicht separat behandelt, sind jedoch in die bereits erwähnten Kategorien integriert. Dies erfolgte um Redundanzen zu vermeiden. Sämtliche Themenkomplexe beinhalten mehrere Unterpunkte, eine detaillierte Übersicht ist im Anhang auf S. 74 zu finden. Alle für ausreichend robust befundenen Evaluierungen wurden dabei zunächst individuell mit Hilfe des Analyserasters Synthese ausgewertet. Im Anschluss daran wurden die Ergebnisse der Auswertung zunächst nach Evaluierungstyp und Themenkomplex gruppiert und zusammengefasst. In einem weiteren Schritt wurde einerseits Evaluierungstyp-übergreifend, andererseits themenübergreifend nach Informationen gesucht. Die so gefundenen Informationen sind in die Unterkapitel zu den einzelnen Themenkomplexen integriert. Auf Basis der Ergebnisse dieser Auswertungen haben wir in einem letzten Schritt basierend auf Erfahrungen in der internationalen Zusammenarbeit, in der internationalen Bildungsentwicklung sowie dem Stand der Forschung, Empfehlungen formuliert.

3.2 Ergebnisse

3.2.1 Manageriale Erfolgs- und Misserfolgskriterien

Ein Ziel war es, die Aussagen in den Evaluierungen den fünf Erfolgsfaktoren von Capacity WORKS zuzuordnen. Dies war jedoch nicht in allen Fällen eindeutig möglich; Probleme bezüglich der Trennschärfe existierten vor allem was die Steuerung, Prozesse und das Lernen betrifft. Wir haben uns deshalb zwar von den Erfolgsfaktoren leiten lassen, sie aber gleichzeitig so interpretiert, dass wir den Aussagen aus den Evaluierungen gerecht werden und diese möglichst authentisch abbilden können.

Aus den vielen einzelnen Aussagen haben wir trotz der erwähnten grundsätzlichen Herausforderungen bezüglich der Vergleichbarkeit Tendenzen über alle Evaluierungstypen hinweg herausgearbeitet. So wurde die Relevanz für die strategische Ausrichtung des Vorhabens mit nur einer Ausnahme (die immerhin zufriedenstellend war) als gut bis sehr gut beurteilt. Den Vorhaben wird attestiert, dass sie sich mit wichtigen Entwicklungshemmnissen befassen und sowohl mit nationalen als auch internationalen Politiken im Einklang sind. Die Kooperation der Regierungspartner wurde hinsichtlich Motivation und Kapazitäten sehr oft bemängelt. Dabei schien aber die Zusammenarbeit mit Regierungspartnern auf der dezentralen Ebene eindeutig besser wegzukommen. Die Zusammenarbeit mit Nichtregierungspartnern kam zwar weniger oft vor, wurde dann aber positiv beurteilt oder die Kritikpunkte bezüglich der Abhängigkeit der NRO vom Vorhaben oder der exklusiven Zusammenarbeit mit nur einer NRO waren vielmehr an die Leitung des Vorhabens adressiert als an die betroffenen NRO selbst. Mit wenigen Ausnahmen stellten die Evaluatoren Schwächen im M+E-System der Vorhaben fest.

Für die UEs wurden in einigen Fällen ein zu vorzeitiges Programmende und eine damit zusammenhängende fehlende Konsolidierung des Erreichten thematisiert. Die PFKs wurden bezüglich strategischer Ausrichtung und Partnerkooperation tendenziell etwas positiver beurteilt als die UEs. Für die DEs war es aufgrund der geringen Anzahl (N=2) kaum möglich Tendenzen abzuleiten.

3.2.1.1 Unabhängige Evaluierungen

Strategische Ausrichtung:

UE	GHA		GUI		HGP		INA		NAM		PAK		PER		CHA		ZA		SV		Ø	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A				
Relevanz	1	ZM	1	ANG	2	ANG	3	ANG	1	ZM	2	ANG	2	ANG	2	ANG	2	ANG	1	ANG	1.7	ANG 8, ZM 2

Tabelle 4: Übersicht der Bewertungen des DAC-Kriteriums Relevanz für UEs⁵

Die Tabelle zeigt, dass die Relevanz für die strategische Ausrichtung der Vorhaben gemäß DAC-Kriterium von den Evaluatoren⁶ als gut bis sehr gut beurteilt wurde. Nur im Falle von Indonesien wurde sie etwas kritischer, aber immer noch als zufriedenstellend bewertet. Den Vorhaben wird attestiert, dass sie auf ein Entwicklungshemmnis eingehen und damit einen Befassungsbedarf de-

⁵ Abkürzungen für die Tabellen 4-11: A – Beurteilung der Angemessenheit der Benotung (durch uns); B – Benotung der DAC-Kriterien durch die Evaluatoren; ANG – angemessenen; ZM – zu mild; ZG – zu streng; Ø B – Durchschnittsbenotung der DAC-Kriterien durch die Evaluatoren. Eine tabellarische Übersicht über die Benotungen aller DAC-Kriterien und der Beurteilung ihrer Angemessenheit findet sich im Anhang auf Seite 162.

⁶ Mit „Evaluatoren“ bezeichnen wir in diesem Dokument nicht uns, sondern die Personen, die die Evaluierungen vorgenommen haben, welche uns zur Auswertung vorgelegt wurden.

cken. Des Weiteren wird oft darauf hingewiesen, dass die Vorhaben der BMZ-Ausrichtung sowie internationalen Richtlinien entsprechen und durch die jeweiligen nationalen Regierungspolitiken des Partnerlandes legitimiert sind. Für Pakistan wird allerdings darauf hingewiesen, dass der Flüchtlingsbereich, mit dem sich das Vorhaben befasst, nicht der BMZ-Ausrichtung entspricht. Die Benotungen des DAC-Kriteriums Relevanz wurden von uns größtenteils als angemessen beurteilt. Eine tabellarische Übersicht über die Benotungen aller DAC-Kriterien und der Beurteilung deren Angemessenheit findet sich im Anhang im Kapitel 6.6 auf Seite 76.

Positive Aussagen zur strategischen Ausrichtung finden sich z.B. für das Vorhaben in Indonesien, dem ein klarer strategischer Ansatz attestiert wird. Im Tschad wird die basisnahe Vorgehensweise bei der Planung gelobt und für Namibia wird gesagt, dass das Anspruchsniveau des Vorhabens insgesamt als realistisch einzustufen ist. Positiv vermerkt wurde für das SV „Innovative Ansätze“ dessen Orientierung an der Education for All – Fast Track Initiative (EFA) und den Millennium Development Goals (MDGs) sowie dessen Fokus auf die Entwicklung und die Anwendung innovativer Ansätze und die systematische Aufarbeitung von Bildungsthemen in der Grundbildung. Das Vorhaben stand damit im Einklang mit den internationalen Vereinbarungen, universellen Zugang zu Bildung zu schaffen und die Bildungsqualität zu erhöhen.

Nebst diesen positiven gibt es auch diverse kritische Anmerkungen zur strategischen Ausrichtung. So wird für einige Vorhaben bemängelt, dass die Zielsetzungen oder – wie im Falle von Ghana – die Maßnahmen zur Risikominderung wenig realistisch seien. Die unrealistischen Zielsetzungen beziehen sich teils auf die Gesamtziele, teils auf einzelne Unterziele, wie im Fall von Namibia auf den muttersprachlichen Unterricht der San. Im speziellen Fall von Pakistan wird der „supply driven“- und der damit zusammenhängende „emergency“-Ansatz in Frage gestellt. Auch wurden dort die Finanzierungsungewissheit und die damit korrespondierende Planungsunsicherheit bemängelt.

Das Programmende wurde in einem Großteil der UEs kritisch thematisiert. Oft waren die politischen Rahmenbedingungen vor Ort und auch seitens der BMZ ursächlich dafür. Im Anhang auf Seite 77 befindet sich ein Überblick über die Programmenden in den einzelnen Vorhaben. Bemängelt wurden in einigen Fällen eine fehlende oder unklare Exit-Strategie sowie eine ungenügende Konsolidierung der Projektaktivitäten. Das Ende wurde zum Teil als zu abrupt oder die Laufzeit als zu kurz empfunden. Im Tschad wurde laut Evaluierung die Entscheidung des BMZ, keiner weiteren Verlängerung zuzustimmen, als Schock und als entmutigend empfunden. Auch in Zentralamerika wurde nach der Fusion zur GIZ eine vom Vorhaben erwartete Konsolidierungsphase nicht bewilligt. Dadurch seien positive Ansätze verpufft. In Peru hätte eine längere Laufzeit gemäß den Evaluatoren eine höhere Nachhaltigkeit bewirkt.

Im Tschad wurde das Fehlen einer Evaluierung gegen Programmende kritisiert. Man war der Meinung, dass eine solche Evaluierung noch Lösungen zur Fortführung des muttersprachlichen Unterrichts und der HIV/AIDS-Aktivitäten gebracht hätte. Im Gegensatz dazu wurde das Vorhaben in Zentralamerika am Ende evaluiert und die Lernergebnisse dokumentiert. Auch in Peru wurden die Programmergebnisse aufgearbeitet.

In Pakistan wurde das Vorhaben in eine lokale NRO umgewandelt, die zum Zeitpunkt der Evaluierung ihre Mandatsbasis erfolgreich ausweiten konnte.

Kooperation mit lokalen Partnern:

Die Kooperation mit lokalen Regierungspartnern auf dezentraler Ebene wurde im Vergleich zu der Kooperation nationaler Regierungspartner positiver bewertet. So wurde im Tschad die enge Zusammenarbeit mit der dezentralen, lokalen Ebene gewürdigt und in Guinea der Einbezug der dezentralen Bildungsebene als gelungen hervorgehoben. In Peru gaben die regionalen Partner im Gegensatz zur nationalen Ebene den Programmenthemen eine hohe Priorität. Die Veränderungsenergie der staatlichen Partner wurde in Peru auf regionaler Ebene als eindeutig höher eingeschätzt. Allerdings sank auf dieser Ebene das Ownership nach einem Regionalregierungswechsel. Als Gegenbeispiele können Indonesien und Namibia aufgeführt werden. In Indonesien hat der fortschreitende Dezentralisierungsprozess zu Defiziten in der Steuerung des Vorhabens geführt und in Namibia war die Unterstützung nationaler Regierungspartner – zumindest am Anfang – sehr hoch.

In dem Vorhaben in Honduras, Guatemala und Peru wurden die privaten Partneruniversitäten als Schlüssel-Akteure hervorgehoben. Sie (und nicht die staatlichen Stellen) waren entscheidend für den Erfolg des Vorhabens. Im Tschad wurde mit bereits existierenden Elternvereinen zusammengearbeitet, die dann auch nach dem Programmende weiter funktional tätig waren.

Die Partner in Peru haben laut Evaluierung den Konsultationsprozess des Vorhabens als sehr positiv empfunden. Ähnliches wird für Namibia berichtet, wo das Vorhaben gemeinsam mit den Partnern geplant und durchgeführt wurde, was zu Ownership auf Partnerseite geführt hat.

In Zentralamerika wurde der durch das Vorhaben ermöglichte Süd-Süd-Austausch von den Partnern als Mehrwert empfunden. Allerdings konnte dieser nicht so etabliert werden, dass er nach dem Programmende fortgeführt worden wäre.

Nebst diesen positiven Erwähnungen gab es in den Evaluierungen auch einige kritische Anmerkungen zu der Kooperation mit den Partnern. So wurden nicht immer alle wichtigen Akteure gebührend in das Vorhaben miteinbezogen, wie etwa im Tschad, in Zentralamerika oder in Namibia. Im Tschad fühlten sich nicht alle Akteure gleich einbezogen. In Zentralamerika war man sich zu wenig bewusst, dass es in einem Sektorreformprozess nebst dem Dialog mit dem Sektorministerium noch andere Ministerien und politische Interessensgruppen zu beachten gilt. In Namibia wurde die Lehrgewerkschaft nicht in das Vorhaben miteinbezogen. Manchmal hing die Durchführung auch von einem motivierten Individuum innerhalb der Bürokratie ab, dies war laut Evaluierung in Ghana der Fall.

In Namibia und dem Tschad wurden eine fehlende Unterstützung der Eltern und der Regierungspartner für den muttersprachlichen Unterricht festgestellt, was auf ein ungenügendes Bewusstsein auf lokaler Ebene in Bezug auf die Vorteile des muttersprachlichen Unterrichts zurückgeführt wurde.

Im Hinblick auf die Zusammenarbeit mit den Regierungspartnern wurden in den Evaluierungen immer wieder limitierende Faktoren aufgezählt; häufig genannt wurden ein fehlender politischer Wille und interne Widerstände. Dabei wurden auch erschwerende Rahmenbedingungen für die Regierungspartner geltend gemacht: umständliche Bürokratie, verschiedene involvierte Akteure,

Arbeitsüberlastung, verzögerte politische Entscheidungen, lange Vakanzen, mangelhafte Ressourcenausstattung oder ungeklärte Fragen der Ressourcenzuteilung. Ein weiteres Thema war die oft fehlende Kontinuität seitens der Regierungspartner, verursacht durch häufige Personalwechsel mit – wie im Falle des Tschads – neuen, als unmotiviert wahrgenommenen Beamten. In den Evaluierungen wird erwähnt, dass die Prioritäten auf Regierungsseite von denen des Vorhabens manchmal abweichen (z.B. in Zentralamerika) oder die Präferenzen bezüglich des Vorgehens nicht immer deckungsgleich sind (z.B. im Tschad). In Pakistan wurde die Zusammenarbeit mit verschiedenen Ministerien mit jeweils unterschiedlichen Vorstellungen als schwierig beschrieben. Bezüglich der Zusammenarbeit mit NRO-Partnern wurde in Pakistan bemängelt, dass dort eine NRO exklusiv unter Vertrag genommen wurde anstatt die Konkurrenz zwischen den verschiedenen NROs zu stimulieren.

Kooperation mit anderen Projekten:

Auch bezüglich der Kooperation mit anderen Projekten ergibt sich ein gemischtes Bild. Positiv herausgehoben wurde im Tschad die enge Zusammenarbeit mit flankierenden GTZ Projekten. In Zentralamerika wurde die Zusammenarbeit mit der UNESCO zur Vermeidung von Doppelungen gelobt. In Zentralamerika wurde allerdings auch eine fehlende Koordinierung und Verzahnung mit anderen EZ-Vorhaben bemängelt. Das SV „Innovative Ansätze“ war von einer Vernetzung mit zahlreichen Partnern gekennzeichnet, was als positiv bewertet wurde. Insbesondere wurde bei Querschnittsthemen mit anderen SV aus den Bereichen Bildung (Konfliktbearbeitung) oder Gesundheit (HIV/AIDS) zusammengearbeitet. Die Zusammenarbeit mit den bilateralen Bildungsvorhaben wurde in diesem SV allerdings als verbesserungswürdig beschrieben.

Projektmonitoring:

Die Evaluierungen monieren mehrheitlich Schwächen im Projektmonitoring der einzelnen Vorhaben. So wurde in Namibia beispielsweise in der Evaluierung das Fehlen eines Monitoringsystems moniert. In Ghana äußerten die Evaluatoren, es sei kein Hinweis auf ein existierendes projektinternes Monitoring-System zu finden gewesen; ferner hätte das Fehlen einer M+E-Dokumentation die (Re-)konstruktion einer Vergleichsgruppe unmöglich gemacht. Auch im SV „Innovative Ansätze“ wurde das Fehlen eines umfassenden Monitoringsystems beanstandet. Zudem wiesen die Wirkungsketten Mängel auf und es lagen keine schriftlichen Dokumente zu den Erfahrungen des SV vor. In Guinea wird das M+E-System als schwach ausgeprägt charakterisiert. Weitere Kommentare lauten: wenig systematisches und wirkungsorientiertes Monitoring (Tschad), nicht alle Projektbeteiligten kennen Ziele und angestrebte Wirkungen (Guinea), späte Einführung eines M+E-Systems (Pakistan) oder es waren keine Dokumente vorhanden, die eine Einsicht in ein projektinternes Monitoring ermöglicht hätten (Pakistan).

Trotz einiger Abstriche positiv wird das Vorhaben in Indonesien bezüglich seines M+E-Systems bewertet. Es wird als ein umfassendes M+E-System beschrieben, das auch auf Schulen angewendet wurde, die nicht vom Vorhaben profitierten. Einzig bemängelt wurde dort das Fehlen von Indikatoren auf der Impact-Ebene. Ferner wurde bemängelt, dass Dokumente aus dem M+E-Bereich Jahre später nicht mehr auffindbar waren, dies wurde jedoch nicht als direkter Mangel des M+E-Systems angesehen.

Effizienz:

UE	GHA		GUI		HGP		INA		NAM		PAK		PER		CHA		ZA		SV		Ø	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A				
Effizienz	2	ANG	2	ANG	3	ANG	2	ANG	4	ANG	2	ANG	3	ANG	2	ZM	2	ANG	4	ZS	2.6	ANG 8, ZM 1, ZS

Tabelle 5: Übersicht der Bewertungen des DAC-Kriteriums Effizienz für UEs

Der Tabelle ist zu entnehmen, dass die Effizienz von den Evaluatoren als weniger gut bewertet wurde als etwa die Relevanz. Im Fall von Namibia wurde sie gar als nicht zufriedenstellend eingestuft. Die Benotungen wurden von uns auf der Basis der vorgefundenen Begründungen bis auf eine Ausnahme (Tschad) als angemessen erachtet. Grundsätzlich muss aber festgehalten werden, dass die Bewertung der Effizienz im Rahmen der Evaluierungen aufgrund fehlender Daten auf der Input- sowie Output-Seite nicht immer vollumfänglich behandelt werden konnte. So wurde beispielsweise im SV „Innovative Ansätze“ festgehalten, dass die Kostendokumentation nicht geeignet war, um Aussagen zur Effizienz zu machen. Sie sei zu allgemein gehalten, als dass die Leistungen des SV den Kosten zugeordnet werden könnten.

Einschränkungen bei der Effizienz wurden zum Teil dort lokalisiert, wo die Wirkungen des Vorhabens im Vergleich zu den eingesetzten Mitteln hinter den Erwartungen zurückblieben und eine Reihe von Outputs nicht erreicht wurden, wie z.B. in Namibia. Auch in Peru führten Einschränkungen hinsichtlich der Nachhaltigkeit zu einer niedriger bewerteten Effizienz. In Pakistan beeinträchtigten die relativ hohen Mitarbeitersaläre und Reisekosten sowie häufige Wechsel der Projektleitung die Effizienz. Auch wurden dort die Kontrollroutinen als mühsam und zeitintensiv bewertet. Im regionalen Vorhaben in Honduras, Guatemala und Peru wurde ein Einsparpotenzial bei den zu langen und zu aufwändigen Fachkursen festgestellt. Positiv bewertet wurde die Effizienz durch die erfolgreiche Zusammenarbeit mit anderen Gebern in Ghana, durch die hohe Nutzung lokaler Ressourcen im Tschad, relativ günstigen Kosten für internationale Mitarbeiter in Pakistan und durch die gute Produktionskosteneffizienz in Peru.

Projektmanagement:

In diesem Kapitel sind weitere Aspekte aufgeführt, die das Projektmanagement betreffen und in den vorherigen Kapiteln noch keine Erwähnung fanden. So wurde der Prozess durch das Vorhaben in Indonesien als effektiv gemanagt beschrieben. Dem gleichen Vorhaben wird auch attestiert, das Lernen und Innovieren gefördert zu haben. In Pakistan wurden die gute Logistik und das gute Organisationsmanagement gewürdigt.

3.2.1.2 Projektfortschrittskontrollen

Strategische Ausrichtung:

PFK	BEN		GUI		HON		COD		MAW		MTN		MOZ		SV		ZAs		Ø	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A				
Relevanz	1	ANG	1	ANG	1	ZM	2	ANG	2	ANG	2	ANG	1	ZM	2	ANG	1	ANG	1.4	ANG 7, ZM 2

Tabelle 6: Übersicht über die Bewertungen des DAC-Kriteriums Relevanz für PFKs

Auch im Rahmen der PFKs wurde die Relevanz der strategischen Ausrichtung als sehr positiv und von uns mehrheitlich als angemessen beurteilt. Im Vergleich zu den unabhängigen Evaluierungen

fällt auf, dass der Programmabbruch nicht thematisiert wird, was insofern logisch erscheint, da es sich bei den PFKs um Evaluierungen von laufenden Vorhaben handelt.

Die strategische Ausrichtung scheint außerdem für die PFKs tendenziell positiver beurteilt zu sein, darauf weisen zumindest die erhöhte Anzahl positiver Aussagen hin. Positiv vermerkt wurde in Mauretanien die flexible Reaktion auf eine veränderte Nachfrage, in Zentralasien das Thematisieren von Risiken und politischer Umsetzbarkeit, in Guinea der Umstand, dass sich die eigene Umsetzung schlussendlich in der Krise als gute strategische Entscheidung herausstellte, in Benin die Ausrichtung auf die Leistungsfähigkeit von Schulämtern als Hebel zur Verbesserung von Dienstleistungen an Schulen, in Honduras die Anpassung der Strategie an den instabilen Kontext und die exemplarische Nutzung von regionalen und subkontinentalen Ressourcen und Potenzialen, in Mosambik die Entwicklung einer Exitstrategie bezogen auf die Einstellung der Zusammenarbeit mit einem einzelnen Regierungspartner und in Malawi die Integration von verschiedenen Ebenen (national, Distrikt, Schule), die es ermöglichte die auf der Schulebene gemachten Erfahrungen in die übergeordneten Ebenen einfließen zu lassen.

Auch für die PFKs wurden einige kritische Punkte aufgeführt. So wurde für Mauretanien bemängelt, dass keine fundierte politökonomische Analyse mit einem Mapping der wichtigsten Akteure, das Interessen, Motivationen und Möglichkeiten der Akteure aufzeigt hätte, gemacht wurde. Es fehle dadurch eine realistische Einschätzung der Reformprozesse. In Honduras wurde beanstandet, dass das Programm nicht spezifisch auf die indigene Bevölkerung eingehe. Zudem wurden in Honduras transversale Themen einbezogen, ohne dass deren Relevanz für den Lehrplan bekannt gewesen wäre. Für den Kongo wurden die Zielsetzungen bezüglich des Stoppens der Landflucht als unrealistisch taxiert. In Benin fehlte eine kohärente Strategie für das Scaling-up. Ähnlich wie das Vorhaben in Pakistan lag das Vorhaben im Kongo nicht im Prioritätsbereich der Zusammenarbeit von Deutschland mit dem Kongo; es fehlte darum ein strategisches Orientierungsdokument. Zudem wurde die Strategie im Kongo mit wenig Einbezug der Partner formuliert.

Kooperation mit lokalen Partnern:

Auch für diesen Aspekt scheinen die positiven Punkte etwas stärker vertreten als bei den UEs. Positiv erwähnt wurde in zwei Fällen – in Malawi und in Zentralasien – die vertrauensvolle, konstruktive und als unterstützend empfundene Zusammenarbeit mit dem Projektteam. In Guinea wurde es von den Partnern außerordentlich geschätzt, dass das Projekt auch während der politischen Krise weiter funktionierte. In Malawi wurde die physische Präsenz der Projektmitarbeitenden mit ihren Arbeitsplätzen im Ministerium von den Partnern geschätzt. In Zentralasien wurden Vorschläge des Vorhabens gemeinsam mit den Partnern diskutiert unter Berücksichtigung der Ressourcenknappheit, was die Wahrscheinlichkeit der Nachhaltigkeit erhöht habe. Es handelte sich um eine Steuerungsstruktur, die eine kontinuierliche Abstimmung mit der Partnerseite beinhaltete. In Benin war das Vorhaben der einzige Entwicklungspartner für die Stärkung der Schulverwaltungen, was diesen Beitrag bei den Partnern besonders erwünscht machte. Das Ausbildungsmodul – von den Evaluatoren als entwicklungspolitischer Leuchtturm bezeichnet – hat laut Evaluierung in Mauretanien viel zum Ansehen des Vorhabens bei den Partnern beigetragen. In Mosambik war die NRO Don Bosco der wichtigste Kooperationspartner und Mittler bei der Umsetzung gemeinsamer Weiterbildungsmaßnahmen. Auch war in Mosambik die Zusammenarbeit auf Zentrumsebene mit der Privatwirtschaft ein Erfolg.

Im Kongo wurde kritisiert, dass das Vorhaben lokale Dienstleistungs-NRO kreiert hat und diese finanziell abhängig vom Projekt waren. In Benin wurde die Kommunikation zwischen dem Vorhaben und den identifizierten Point focaux (Ansprechpartnern) sowie dem hochrangig besetzten Steuerungskomitee zu wenig strategisch genutzt. In Mosambik wurde mangelnder Umsetzungswille und Beratungsresistenz bei einem staatlichen Partner festgestellt, mit dem die Kooperation dann allerdings abgebrochen wurde. In Honduras wurden nicht alle Verbündeten miteinbezogen und das Projekt im Kongo hat laut Evaluierung die Zusammenarbeit mit der Provinzregierung gar nicht erst gesucht. Dabei wird angemerkt, dass es fraglich sei, wie wichtig der Sektor für die Regierung war, da diese für den Sektor keine finanziellen Mittel budgetiert hatte.

Kooperation mit anderen Projekten:

In Mauretanien wurde die gute Abstimmung mit dem UNDP hervorgehoben. Ansonsten überwogen die kritischen Kommentare: In Honduras war die Geberkoordination nicht effektiv, was zu Duplikationen und parallelen Unterstützungen führte. Auch funktionierte das subkontinentale Netzwerk nicht, da die GIZ nur noch ein Bildungsprojekt in der Region, in Guatemala unterhielt. Eine Kooperation mit anderen Projekten wurde im Kongo gar nicht abgeklärt. Auch in Benin fand keine systematische Kooperation mit anderen Gebern statt, was jedoch auch daran lag, dass die GIZ das einzige Vorhaben war, welches direkt auf Ebene der Schulämter tätig war. In Malawi führte eine fehlende Abstimmung der Akteure der zahlreichen Workshops, organisiert durch die verschiedenen Projekte, zu vielen Absenzen seitens der Regierungspartner.

Projektmonitoring:

Wie bei den UEs überwogen auch bei den PFKs die kritischen Kommentare: In Guinea existierte ein Monitoringsystem, es wurde jedoch nicht systematisch genutzt; in Honduras war das institutionalisierte Gedächtnis des Programms verbesserungswürdig; in Benin war das M+E-System zwar detailliert, doch zu wenig auf Wirkung ausgerichtet; in Sektorvorhaben schien das Wissensmanagement ausbaufähig; im Kongo wurde das Impact-Monitoring nicht systematisch angewendet; in Malawi fehlte es den Lernprozessen an Komplexität und Diversität, es fehlten Informationen, die die Wirkung auf der Ebene der Schule abbilden und das Team hatte nur eine vage Vorstellung von Monitoring.

Positiv vermerkt wurden Mauretanien und Mosambik. In Mauretanien wurde dem Vorhaben ein gut entwickeltes Aktivitätsmonitoring attestiert. In Mosambik wurde ein detailliertes, komplexes und aufwändiges Monitoring eingeführt. Die Ergebnisse werden dort dreimal im Jahr erhoben, präsentiert, diskutiert und handlungsorientiert dokumentiert.

Effizienz:

PFK	BEN		GUI		HON		COD		MAW		MTN		MOZ		SV		ZAs		∅	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A				
Effizienz	3	ANG	2	ANG	2	ANG	4	ANG	4	ZS	3	ANG	2	ANG	3	ANG	3	ANG	2.9	ANG 8, ZS 1

Tabelle 7: Übersicht über die Bewertungen des DAC-Kriteriums Effizienz für PFKs

Wie für die UEs gilt auch für die PFKs, dass es für die Evaluatoren nicht einfach war, die Effizienz aufgrund der vorhandenen und von den Vorhaben zur Verfügung gestellten Informationen klar zu beurteilen. Tendenziell wurde die Effizienz, wie obige Tabelle zeigt, ähnlich bewertet wie für die UEs und auch von uns als angemessen beurteilt.

In Honduras wurde positiv hervorgehoben, dass das Vorhaben mit wenig internationalen externen Experten und Beratern arbeitete, was Reisekosten reduzierte und die Nutzung der Kapazitäten der eigenen Equipe optimierte. In Zentralamerika wurde die Effizienz durch eine enge Abstimmung mit anderen Gebern und Forschungseinrichtungen erhöht. Es wurden dort auch etliche Leistungen durch Mitarbeiter des Bildungsministeriums erbracht.

Durch das schwierige Umfeld im Kongo blieb die Effizienz des Vorhabens hinter dem zurück, was sonst erwartet werden konnte. In Benin wurden die hohen Personalausgaben, die vielen hochpreisigen Workshops und die fehlende Hebelung von Korbmitteln kritisiert.

Projektmanagement:

Die Zusammenarbeit im Team wurde im Sektorvorhaben von den Teammitgliedern als nicht zufriedenstellend empfunden - das Team bestehe aus Einzelkämpfern und die Aufgabenverteilung sei unklar. In Zentralamerika gab es keine gemeinsame Jahresplanung und keinen Austausch zwischen den drei Projektbüros. Auch fehlte ein übergreifender Einsatzplan der Kurzzeiteinsätze. In Benin fehlten ein systematisches Prozessmanagement, das Suchen und Einbringen von Innovationen sowie die nötige Kontinuität von internationalen Experten im Team für die Einführung von Capacity WORKS. Das Team in Malawi kritisierte die zu knappe Zeit für Austausch und Koordination.

3.2.1.3 Dezentrale Evaluierungen

Strategische Ausrichtung:

In den zwei Evaluierungen war nur wenig zur strategischen Ausrichtung zu finden. In Zentralamerika wurde dem Vorhaben eine zielgruppen- und bedarfsorientierte Planung attestiert. Im Kosovo wurden die Indikatoren als sehr ambitioniert eingestuft, die den Effekt hatten Ressourcen zu mobilisieren.

Kooperation mit lokalen Partnern:

In Zentralamerika wurde von den Partnern eine große Wertschätzung gegenüber der Projektleitung ausgedrückt. Durch einen partizipativ gestalteten Prozess konnte das Ownership bei den Partnern gesichert werden. Die Mitarbeiter des Paulo Freire Instituts und der Universität Javeriana konnten die Philosophie von EduPAZ authentisch weitergeben. Im Kosovo wurden die Meetings mit verschiedenen NROs als einmaliger Raum für das Networking wertgeschätzt.

Kritisiert wurde im Kosovo die fehlende Koordination zwischen den verschiedenen Bemühungen im Jugendsektor. In Zentralamerika stellt sich die Entwicklung eines Diplomkurses aufgrund des partnerschaftlichen Ansatzes als sehr zeitaufwändiger Prozess heraus. In der zweiten Phase hat die GIZ die Teilnehmenden für die Kurse in Deutschland eigenhändig aus den staatlichen Kooperationspartnern gewählt, ohne Ausschreibungsverfahren. Dadurch nahmen viele Beamte und kaum mehr NRO-Vertreter teil, was der Dynamik Abbruch tat.

Übrige Punkte:

Für Zentralamerika wurde erwähnt, dass Schnittstellen zu anderen GIZ Projekten gesucht wurden. Dort wurde auch ein umfassender M+E-Prozess in der Pilotphase angewendet und zudem wurden

die Feedbackbögen der Teilnehmenden ausgewertet und für Anpassungen genutzt. Bemängelt wurde, dass die Arbeitsflüsse durch häufige Wechsel der Programmverantwortlichen unterbrochen wurden. Für den Kosovo gibt es keine besonderen Erwähnungen zu den übrigen Punkten.

3.2.2 Sektorspezifische bzw. fachliche Erfolgs- und Misserfolgskfaktoren

Die Auswertung der Evaluierungen zeigt deutlich, dass nicht alleinstehende Faktoren einen sektorspezifischen bzw. fachlichen Erfolg oder Misserfolg begründen. Auch sehr vielversprechende Ansätze entfalten meist erst im Zusammenspiel mit anderen Maßnahmen ihre Wirkung. Dies wird im Folgenden anhand von Themen aufgezeigt, die in jeweils mehreren Evaluierungen angesprochen wurden.

3.2.2.1 Unabhängige Evaluierungen

Stärkung der Lehreraus- und -fortbildung:

Die Verbesserung von Bildungszugang und -qualität bedingt u.a. eine zahlenmäßig ausreichende und qualifizierte Lehrerschaft, beides ist in den meisten Ländern der Vorhaben nicht ausreichend gegeben. Entsprechend werden in allen Grundbildungsvorhaben in unterschiedlichen Formen und Ausprägungen Lehrkräfte aus- und fortgebildet.

Vor allem in Subsahara-Afrika stellt der deutlich verbesserte Bildungszugang eine enorme Herausforderung für das Bildungssystem dar. Um dem Lehrermangel zu begegnen, wurde z.B. im Tschad die Ausbildungszeit für Lehrer um ein Jahr reduziert. Trotzdem beträgt die Lehrer-Schüler-Relation 1:68.

In Pakistan wurden ca. 7000 Lehrkräfte in Methodik und Didaktik, schülerzentriertem Unterricht und im Gebrauch des im Rahmen des Vorhabens entwickelten Unterrichtsmaterials ausgebildet. Ziel war es vor allen Dingen den afghanischen Flüchtlingen eine bessere Bildungsperspektive zu bieten. Es wurde erwartet, dass die Aus- und Weiterbildung von Lehrkräften in Kombination mit der Verteilung von inhaltlich besserem Unterrichtsmaterial einen stärker schülerzentrierten Unterricht bewirken würde, welcher zu mehr Aktivität, Kreativität und Initiative bei den afghanischen Schülern führen würde. Dieser Effekt konnte nicht nachgewiesen werden. Die Sprechzeit von Lehrern der vom Vorhaben geförderten Schulen blieb mit 62% ähnlich hoch wie diejenige in anderen Schulen mit 64%. Es blieb dabei, dass meist geschlossene Fragen gestellt wurden, die von den Schülern Auswendiglernen verlangen. Die Lehrkräfte brachten mit einem durchschnittlichen Bildungsniveau auf Sekundarstufe, ohne Lehrerbildung auf Tertiärstufe, nicht die Voraussetzungen mit, als dass die durchgeführten Schulungen genügend wirksam werden konnten. Zudem behinderten Rahmenbedingungen wie eine schlechte Infrastruktur und zu wenig Unterrichtsmaterial in oft schlechter Druckqualität die Umsetzung eines stärker schülerzentrierten Unterrichts. Allerdings konnte der afghanischen Regierung eine Liste mit den Namen aller durch das Vorhaben ausgebildeter Lehrkräfte, Schulleiter, Bereichsleiter und Ausbilder übergeben werden, in der Annahme, diese Liste unterstütze die Entwicklung der Grundschulbildung in Afghanistan, da so die Möglichkeit gegeben war, auf diese Personen in Afghanistan zurückzugreifen.

Die Home Schools in Pakistan und die Gemeindelehrer im Tschad sind Beispiele dafür, dass unter gewissen Umständen auch gering oder nicht qualifizierte Lehrer wirkungsvoll unterrichten können. In den Home Schools unterrichten meist Mütter, die dafür weitergebildet wurden. Der Unterricht

findet zu Hause statt. Der Lernstoff der sechs Primarschulklassen, der gleiche wie in formalen Grundschulen der Flüchtlingsdörfer, wird in 30 Monaten (5 Monate pro Schuljahr) vermittelt. Die Stundenpläne sind flexibler, pro Klasse werden durchschnittlich 25 Kinder in ihrer jeweiligen Nationalsprache unterrichtet. Schüler in Home Schools erzielen durchwegs signifikant bessere Leistungen als Schüler in den anderen Schulen. Als ein Grund wird das informelle und dadurch schülerfreundlichere Setting gesehen, das speziell für Mädchen förderlich ist. Im Tschad erzielten Schüler von Gemeindelehrern in einem Vergleichstest bessere Leistungen und das obwohl die Hälfte der Gemeindelehrer für ihre Lehrtätigkeit nicht qualifiziert war und sie wesentlich weniger verdienen als staatliche Lehrer. Dies wurde mit dem persönlichen Engagement und der Motivation der Gemeindelehrer erklärt, die von der Gemeinde beauftragt und beaufsichtigt werden.

In den Vorhaben in Honduras, Guatemala, Peru wurde u.a. bei der Weiterbildung der Lehrerbildner angesetzt, um so die Ebene der Grundschullehrer zu erreichen. Es wurden neue Erkenntnisse aus der Erziehungswissenschaft und der Fachdidaktik für Naturwissenschaften vermittelt um die pädagogische Professionalität der Teilnehmer zu stärken. Auf individueller Ebene wurde ein Nutzen festgestellt, aber kaum Wirkung auf das Bildungssystem. Ein Problem, das nicht gelöst wurde, stellt die vielfach unzureichende pädagogische Arbeit von Lehrern dar. Ihnen fehlen oft sowohl fachliche Kompetenzen zum Unterrichten auf Primarschulebene als auch interkulturelle Kompetenzen, um auf soziokulturelle Besonderheiten der Schüler einzugehen. Hier konstatiert die Evaluierung eine zu ehrgeizig formulierte Wirkungskette, da für die erwarteten Wirkungen auf der Ebene der Lehrkräfte vielfältige Faktoren wie z.B. der Arbeitsmarkt und Arbeitsbedingungen für Lehrer mitspielen. Es wird vermutet, dass das Ziel höchstens langfristig über einen Generationenwechsel erreichbar ist.

In Indonesien wurden zur Festigung der Weiterbildungen in schülerzentrierten Methoden für den naturwissenschaftlichen Unterricht und deren Verwendung Lehrer-Arbeitsgruppen gebildet. Diese wurden nach Ende des Vorhabens weitergeführt, allerdings zur Diskussion verschiedenster Anliegen genutzt.

In Namibia basiert ein Schulbuch für den Alphabetisierungsunterricht auf neuen, international erprobten Lernkonzepten für die erste Klasse. Diese waren aber einem Großteil der Lehrkräfte in Namibia nicht bekannt. Hier fehlten entsprechende Fortbildungsangebote in der Unterrichtung afrikanischer Muttersprachen sowie in schülerorientierten Unterrichtsmethoden.

In Guinea bewirkten Qualifizierungsmaßnahmen für Lehrkräfte und Schulleiter, dass die Lehrkräfte anschauliche Unterrichtsmaterialien selber herstellen, dass sie den Unterricht besser vorbereiten und dass Unterricht in kleinen Arbeitsgruppen organisiert wird. Auch bezogen die in „Gleichberechtigung der Geschlechter“ ausgebildeten Lehrkräfte die schüchternen Mädchen besser in den Unterricht ein. Dies führte zu einer angstfreieren Beteiligung der Schüler am Unterricht und trug zu besseren Leistungen in Mathematik und Französisch bei.

Entwicklung, Verbreitung und Einführung von Lehrmitteln und Unterrichtsmaterial:

Die Entwicklung, Verteilung und Einführung von Lehrmitteln und Unterrichtsmaterial auf Ebene der Lehrerbildung und/oder auf Schulebene wurde in mehreren Vorhaben als ein wichtiges Instrument zur Förderung der Bildungsqualität eingesetzt. Die Evaluierungen zeigen, dass ein Zusammenspiel mehrerer Faktoren nötig ist, damit sich dieses Potenzial wirkungsvoll entfalten kann.

Ideal war das Vorgehen bei der Entwicklung von Lehrmitteln in Ghana. In verschiedenen ghanaischen Sprachen wurden innovative Lehrmittel in sehr guter Qualität entwickelt. Sie nehmen Bezug zur Lebenswelt der Schüler, basieren auf aktiven Lehrmethoden und ermöglichen schülerzentrierten Unterricht, sind gendergerecht konzipiert und zeigen positive Werte und Gewohnheiten auf. Sie sind in einfacher Sprache gehalten und attraktiv illustriert und gestaltet. Zusätzlich zu den Lehrmitteln wurden Lehrerhandbücher entwickelt, welche Lehrkräften helfen den Aufwand für die Unterrichtsvorbereitung zu vermindern. Das Material wurde unter Einbezug des Personals des Bildungsdepartements auf Distriktebene sowie der Lehrerbildungsinstitutionen entwickelt. Damit wurden die Entwicklungsarbeiten abgestützt und bei diesen Schlüsselpersonen konnte relevantes Wissen aufgebaut werden. Auf individueller Ebene wurde damit erfolgreich Kompetenz aufgebaut, aber es wurden zu wenige Personen erreicht und u.a. dadurch eine Systemwirkung verpasst. Hinderlich für die Verbreitung erwies sich auch, dass das Material nur für die Unterstufe und nicht für die Klassen vier bis sechs entwickelt wurde und die Übersetzung in weitere sechs Sprachen fehlte. Die neuen Lehrmittel wurden zu kurze Zeit und nur auf Ebene der Lehrerbildung eingesetzt, daher konnte keine Wirkung auf Schülerebene erreicht werden.

In Guinea wurde anschauliches pädagogisches Unterrichtsmaterial zur Unterstützung von schülerzentriertem Unterricht im Rahmen der Ausbildung von Primarlehrern und in Indonesien „Science Kits“ für Schüler/-innen für einen qualitativ verbesserten naturwissenschaftlichen Unterricht im Rahmen von Lehrerweiterbildungen hergestellt. Beide Male wurde lokal verfügbares Material benutzt, was die Herstellungskosten gering hielt und eine einfache eigene Reproduktion ermöglichte. Die Verbindung mit Lehreraus- oder -weiterbildung sicherte eine Verankerung der Verwendungweise des neuen Materials und unterstützte zusätzlich dessen langfristige Verwendung. Im Vorhaben Indonesien wurden – im Gegensatz der Kits für die Schüler/-innen – für Lehrkräfte besondere „Science Kits“ in Deutschland hergestellt. Zwar wurden diese hochwertig produziert, aber kaputte oder fehlende Teile können kaum ersetzt werden. Die Ersatzteile sind lokal nicht erhältlich und eine Bestellung beim Lieferanten in Deutschland ist administrativ aufwändig und teuer. In Indonesien wurden nur 2-3% aller Primarlehrkräfte im Gebrauch der „Science Kits“ weitergebildet und diese Weiterbildungen wurden nach Ende des Vorhabens nicht fortgeführt. Dadurch wurden eine breite Einführung und eine Systemwirksamkeit nicht möglich.

Vom Vorhaben in Namibia entwickelte Lehr- und Lernmaterialien wurden nicht in Pilot Schulen auf lokaler Ebene getestet, was eine Überarbeitung und Qualitätssicherung ermöglicht hätte. Es wurden qualitative Mängel wie eine fehlende chronologische Ausrichtung der Bücher auf die neuen Lehrpläne, das Fehlen eines Arbeitsbuchs für Mathematik und ein zu hohes Anspruchsniveau für den Lese- und Schreibunterricht der ersten Klasse bemängelt.

In Pakistan wurden mehr als 135 Unterrichtsmaterialien in Dari und Pashto entwickelt, gedruckt und verteilt. Es handelt sich dabei meist um Leitfäden für Lehrkräfte und zusätzliches Unterrichtsmaterial, aber auch Studien, Bücher und wissenschaftliche Berichte. Diese Materialien wurden in großer Zahl von anderen lokalen und internationalen Institutionen in Pakistan und Afghanistan verwendet.

Regionaler Dialog und Austausch:

In mehreren Vorhaben wurden lokale und regionale Austauschmöglichkeiten geschaffen, um dadurch die Nutzung neuer Methoden und Ansätze zu festigen und zu verbreiten. In Zentralameri-

ka konnte der regionale Dialog in Verbindung mit einem Blended Learning Modul Entscheidungsträger sensibilisieren und einen Beitrag zur Formulierung von Reforminitiativen leisten. Die Umsetzung erfolgte jedoch nicht. Außerdem wurde ein Süd-Süd-Austausch durchgeführt, welcher den Austausch von erfolgreichen Praktiken in ähnlichem Umfeld ermöglichte. Dieser wurde aber nicht verankert und nur punktuell weitergeführt. Zudem fehlte Zeit um die Effekte zu konsolidieren.

In Peru wurde die Internetplattform Ciberdocencia mit Informationen und Inhalten rund um das Thema Lehrerfortbildung insbesondere von der direkten Zielgruppe der Lehrer intensiv genutzt. Jährlich wurden bis zu 0,5 Mio. Zugriffe gezählt. Die Plattform wurde nach Projektende noch eine gewisse Zeit eigenständig aufrechterhalten. Sie besteht nach wie vor, wurde aber nicht mehr gepflegt.

Das Vorhaben in Namibia unterstützte das Schulverbundsystem. Besonders ländliche Schulen profitierten vom gegenseitigen Austausch, von der gemeinsamen Nutzung der Infrastruktur der Mittelpunktschulen und von der Erhöhung der Prüfungsstandards über die Ausarbeitung gemeinsamer Prüfungsfragen und Bewertungsbögen.

Muttersprachlicher Unterricht:

Die Einführung und Verbreitung von muttersprachlichem Unterricht in afrikanischen Sprachen wurde in mehreren Bildungsvorhaben in Afrika gefördert. Muttersprachlicher Unterricht hat nachweislichen Einfluss auf die Verbesserung der Bildungsqualität⁷. Trotzdem trifft dieser Ansatz in afrikanischen Ländern häufig auf Widerstände, weil das Wissen um die positive Wirkung von muttersprachlichem Unterricht fehlt und die Kolonialsprachen als Unterrichtssprachen bevorzugt werden oder auch weil politisches Kalkül im Spiel ist. Letzteres war z.B. in Guinea der Fall. Als Ausdruck der Abgrenzung zur vorangegangenen Regierung Sekou Tourés wurde die Nutzung der Muttersprache als Unterrichtssprache zurückgenommen und vor allem in der offiziellen Landessprache Französisch unterrichtet.

In Namibia wurde zu sieben afrikanischen Sprachen je ein Sprachenkomitee gegründet; diese Komitees entwickelten nicht nur Lehr- und Lernmaterial für den muttersprachlichen Unterricht, sondern betrieben in vielen Fällen Grundlagenforschung für diese Sprachen, z.B. hinsichtlich der Regeln der Verschriftlichung oder der Standardisierung für die Übersetzung von Fachvokabular. Die Nutzung und der Stellenwert der sieben afrikanischen Sprachen, in welchen das Vorhaben arbeitete, wurden gestärkt. Den erarbeiteten Materialien wird hohe Qualität attestiert. Der Zugang zu qualitativ hochwertigen Lehr- und Lernmaterialien wurde dadurch stark verbessert. Kombiniert wurde dies mit Beiträgen des Vorhabens zur Ausarbeitung von Lesematerialien sowie zur Straffung des Curriculums und der Lehrpläne der unteren Primarstufe. Leider wurden eigentlich geplante Lehrerfortbildungen zur Nutzung der Schulbücher und zur Einführung schülerorientierter Unterrichtsmethoden, auf welchen die neuen Unterrichtsmaterialien basierten, nicht durchgeführt. Auch lagen die Fächerrichtlinien nicht in den afrikanischen Sprachen vor, was die curriculum-konforme Unterrichtsvorbereitung in den afrikanischen Sprachen erschwerte. Die Sprachenkomitees waren lediglich ad-hoc Arbeitsgruppen und nicht in die Struktur des nationalen Curriculuminstituts veran-

⁷ Vgl. K. Komarek: www.bev-bf.de/texte/Sprachen_Komarek.pdf

kert. Nach Ende des Vorhabens wurden sie aufgelöst und die Produktion der ergänzenden Lehrmaterialien eingestellt. Auch wurde die nationale Sprachenpolitik nicht verabschiedet. Schwer wog zudem der fehlende Rückhalt auf der Ebene der Eltern und Lehrer, welche den Unterricht in englischer Sprache forderten. Ab 2005 fehlte die Fortführung von Öffentlichkeitskampagnen zur Stärkung der Akzeptanz der afrikanischen Sprachen als Unterrichtssprachen. Die Nutzung von afrikanischen Sprachen als Unterrichtssprache in den ersten drei Klassen ging in der Folge wieder stark zurück.

Im Tschad wurde neben weiteren pädagogischen Ansätzen der muttersprachliche Unterricht erfolgreich eingeführt. Vergleichsstudien während der Projektlaufzeit belegen, dass Schüler, die vom muttersprachlichen Unterricht und von den durchgeführten Sommerkursen profitierten, bessere Schulleistungen erbrachten als jene, die nicht davon profitierten. Es erfolgte jedoch keine Übernahme ins System und nach Ende des Vorhabens wurde der Ansatz nicht fortgeführt. Dafür werden unterschiedliche Faktoren verantwortlich gemacht: Es fehlte eine institutionelle Verankerung und eine kompetente Struktur, welche die pädagogischen Ansätze weiter förderte. Einige Mitglieder von Elternvereinen, Lehrkräfte und Personen aus den Schulbehörden waren gegen den muttersprachlichen Unterricht und forderten die Rückkehr zu Französisch oder die Einführung von klassischem Arabisch als Unterrichtssprache. Nicht alle Projektbeteiligten waren ausreichend überzeugt davon, dass muttersprachlicher Unterricht auch positive Auswirkungen auf das Lernen der Verkehrssprachen Französisch und Arabisch hat. Es wird festgehalten, dass die Ziel- sowie weitere Stakeholdergruppen über diesen Ansatz und seine Effektivität nicht hinreichend informiert waren und dass eine bessere Sensibilisierung vor der Umsetzung die Nachhaltigkeit dieses Ansatzes gestärkt hätte. Es fehlte auch eine Evaluierung der Wirkung des Ansatzes zu Projektende und damit evidenzbasierte Wirkungsnachweise. Ohne solche war die Bereitschaft zur Fortführung und Ausweitung des muttersprachlichen Unterrichts bei den Entscheidungsträgern nicht gegeben. Das Risiko der Nicht-Akzeptanz durch einzelne Beteiligte war zwar schon zu Beginn des Vorhabens bekannt. Es wurde aber als gering eingeschätzt, weil die Verwendung der Nationalsprachen im Unterricht in der tschadischen Verfassung verankert und damit formal nicht angreifbar ist. Es wird vermutet, dass deshalb keine spezifischen risikomindernden Maßnahmen ergriffen wurden.

Die ghanaische Sprachenpolitik sieht vor, dass der Unterricht in den Klassen 1 - 3 in den ghanaischen Sprachen gehalten wird. Dies wurde 2002 unterbrochen, als der Bildungsminister Englisch als Unterrichtssprache bestimmte. Darauf kam es zu breiten Diskussionen über die Verwendung der Muttersprache in den ersten Bildungsjahren auch für das Lernen einer Zweit- und Drittsprache, an denen sich das Vorhaben intensiv und erfolgreich beteiligte. Nach landesweiten Protesten wurde die alte Sprachenpolitik 2004 erneut offiziell verabschiedet. Vom Vorhaben wurden Lehr- und Lernmaterialien in guter Qualität für Lesen, Schreiben und Naturwissenschaften in den 11 wichtigsten Sprachen in Ghana entwickelt, womit potenziell 85% der Schüler in Ghana erreicht werden konnten. Unter anderem aufgrund personeller Wechsel sowie politischer Veränderungen waren die Bestrebungen, die Methoden und Materialien in die Lehrpläne der Lehrerausbildungsinstitute und Schullehrpläne zu integrieren, nicht erfolgreich. Die Evaluatoren bemängeln in diesem Zusammenhang auch, dass die Makro-, Meso- und Mikroebene nicht systematisch genug in die Implementierungsstrategie aufgenommen wurde. Nichtsdestotrotz wurden laut der Evaluatoren direkt im Material und in den Fortbildungen in ghanaischen Sprachen curriculare Veränderungen eingeführt.

Innovative Ansätze:

Die Home Schools in Pakistan und das Programm FIERE (Filles éduquées réussissent) in Guinea sind innovative Ansätze, mit denen sehr erfolgreich neue Wege gegangen wurden. In Pakistan wurde mit den Home Schools ein informelles Bildungsangebot eingerichtet. Damit wurden Kinder erreicht, die aufgrund ihres Alters nicht mehr in die formelle Schule integriert werden konnten. Alle Schüler absolvierten standardisierte Prüfungen, die ihnen den Übertritt in formale Schulen ermöglichen. Speziell Mädchen im Pubertätsalter, die von ihren Eltern nicht in die Schule geschickt wurden, konnten so auch in den Genuss von Bildung kommen. Fast 43'000 Schüler schlossen die Home School ab, davon 45% Mädchen.

Das Programm FIERE richtet sich an Mädchen der 5. und 6. Klasse, die zum Teil aus besonders benachteiligten Familien stammen. Angeboten wird gezielter Nachhilfeunterricht in Französisch und Mathematik, bei dem auf die Bedürfnisse und Schwierigkeiten der Mädchen eingegangen wird. Kleine Klassen werden von einem Tandem motivierter und gut ausgebildeter Lehrerinnen auf anschauliche Art und Weise unterrichtet. Von Beginn an werden die Bevölkerung, insbesondere die Mütter, und die dezentralen Bildungsstrukturen einbezogen. Aufklärung über HIV/AIDS und weitere sexuell übertragbare Krankheiten, die üblicherweise im kulturellen Kontext in Guinea tabu sind, konnte in den Unterricht integriert werden. Dies gelang in der Zusammenarbeit mit den Ältestenräten, religiösen Führern und Elternvertretern, die das Programm genehmigten und bescheinigten, dass es nicht im Widerspruch zum Koran steht. Dabei wurde der Begriff Sexualerziehung durch den der Verhaltenserziehung ersetzt. Eine Tracer-Studie von 2012 zeigte ein sehr positives Ergebnis bezüglich Verbleibs- und Abschlussquoten. 85% der 9210 FIERE-Mädchen besuchten die Sekundarschule oder das Gymnasium, von ihnen haben 79% mindestens einen durchschnittlichen Notendurchschnitt erreicht, gegenüber 52% bei nicht-FIERE-Mädchen. Es konnten auch Verhaltensänderungen bezüglich Bildung bei den Mädchen und ihren Eltern festgestellt werden: Die Mädchen sind selbstbewusster geworden und möchten im Schulsystem verbleiben. Und die Eltern geben den Mädchen mehr Freiraum zum Lernen und weniger häusliche Pflichtaufgaben.

3.2.2.2 Projektfortschrittskontrollen

Wie in den UEs werden in den PFKs die Qualifizierung von Personal im Bildungssektor, die Entwicklung und Verbreitung von Lehrmitteln und Unterrichtsmaterial, die Einführung effizienterer Unterrichtsmethoden und Maßnahmen zur Förderung von regionalem Dialog und Austausch als wichtige Zugänge genannt, um auf eine Verbesserung von Zugang und Qualität der Bildung hinzuwirken. Es zeigen sich aber weitere Kombinationen und auch zusätzliche Maßnahmen und Aspekte, die wir nachfolgend aufzeigen.

Das Vorhaben in Zentralasien fokussierte ursprünglich sowohl auf die Verbesserung der Bildungsqualität als auch des Bildungsmanagements. Weil jedoch Maßnahmen im Bereich des Bildungsmanagements zu tief in die nationalen politischen Strukturen der Partnerländer einzugreifen drohten, beschränkte sich das Vorhaben im Projektverlauf aus Machbarkeitsgründen auf die Bildungsqualität. Dabei setzt das Vorhaben parallel auf der Mikro-, Meso- und Makroebene an. In den Lehrerbildungsinstitutionen der Partnerländer wurden Innovationen mit hohem Potenzial für eine nachhaltige Verbesserung der Relevanz der Bildung und der Verwertbarkeit schulischer Lerninhalte eingeführt. Dabei geht es insbesondere um naturwissenschaftliche Unterrichtsmethoden mit schülerzentrierten Ansätzen und berufsorientierten Technologieunterricht. Pilotansätze wurden nachfrageorientiert konzipiert, sind replizierbar und an das sozio-kulturelle und politische Umfeld

angepasst. Politischer Rückhalt für die breitenwirksame Umsetzung wird von Schlüsselakteuren auf der Partnerseite zugesichert und von qualifizierten Institutionen getragen, die durch Beratung und Kapazitätsaufbau durch das Programm in diesem Prozess unterstützt werden. Zum Teil gehen schon jetzt vom Vorhaben erprobte Modelle, Konzepte und Verfahrensweisen in die Umsetzung der nationalen Strategien und Reformen ein. Häufige Kurzeiteinsätze und die Ausstattung von Praxisschulen waren zwar teuer, haben aber qualitativ hochstehende Innovationen ermöglicht, die gut aufgenommen wurden. Auf der Mesoebene setzte die Neuformulierung von Lehrplänen für die Lehrerbildung an. Die Vorschläge wurden gemeinsam mit den Partnern diskutiert, um Nachhaltigkeit zu erzielen. Auf der Makroebene wurde ein regionaler Bildungsdialog erfolgreich etabliert. Dabei wurde die politische Umsetzbarkeit berücksichtigt. So ist im Fall von Usbekistan die Durchführungsebene und nicht die Ebene des Bildungsministeriums beteiligt. Die verschiedenen Innovationen wurden aber nicht mit einem systematischen Wirkungsmonitoring erfasst und wenig koordiniert. Um Breitenwirksamkeit und Nachhaltigkeit zu erreichen, wird dem Vorhaben empfohlen, ein Wirkungsmonitoring-System möglichst bei den Partnern zu entwickeln und eine Verbreitungsstrategie inklusive eines Finanzierungsmodells auszuarbeiten.

Das Grundbildungsvorhaben in Guinea setzt die Zusammenarbeit im Rahmen von Vorgängerprojekten fort. 9.5% aller Lehrkräfte wurden in Französisch, Mathematik und Methodik-Didaktik qualifiziert. Auch wenn weitergebildete Lehrer die Programmregion verlassen, sind ihre Kapazitäten dem Land nützlich. Es wurde besserer Unterricht direkt in den Klassen oder in Vorbereitungsdokumenten beobachtet. Lehrkräfte zeigen auch weniger Widerstand in einer 5. oder 6. Klasse zu unterrichten. Das ist bislang bei Lehrkräften mit geringer Bildung häufig der Fall, weil sie auf dieser Schulstufe Mühe bekunden. Einige Lehrkräfte haben Beförderungstests bestanden. Eine Wirkung auf bessere Schülerleistungen wird vermutet, ist aber bisher nicht nachgewiesen. Durch das Vorhaben qualifizierte Schulleiter haben die Schuladministration sichtbar verbessert. Sie haben ein Schulumanagement eingeführt und halten die Verwaltungsinstrumente auf dem Laufenden. Das Vorhaben ist in den dezentralen Strukturen des Bildungssystems stark verankert. Es setzt damit nicht nur an der Basis an sondern auch in der Struktur und erzielt damit Effizienz. Guinea befindet sich in einem grundsätzlichen Prozess der sozio-politischen Transition. Das Vorhaben trägt zu einer positiven Öffnung von Jugendlichen und Erwachsenen, zur Modernisierung des Bildungssystems und zur Emanzipierung der Akteure bei. Eine zentrale Herausforderung besteht in der Finanzierung, da das Programm die aktuellen finanziellen Möglichkeiten des Landes übersteigt und eine zusätzliche Finanzierung nötig bleibt.

In Malawi wird der Dezentralisierungsprozess durch Beiträge zur Gestaltung und Umsetzung von Bildungsplänen der Distrikte unterstützt. Auf der Ebene des Bildungsministeriums unterstützt das Vorhaben die Kapazitätsentwicklungsstrategie durch die Ausweitung des Bildungs-Informationssystems und Qualifizierungsmaßnahmen für Planer auf Divisionsebene. Auf Ebene der Lehrerbildungsinstitutionen setzt das Vorhaben mit Managementberatung und einer Veränderung der Lehrerausbildung in Richtung schülerzentrierter Ansätze an. Lehrer und Schulleiter werden insbesondere in schülerzentrierten Unterrichtsansätzen qualifiziert. Es fehlen allerdings Informationen darüber, ob Wirkungen auf der Ebene der Schülerleistungen eintreten. Deshalb empfehlen die Evaluatoren eine vermehrte Wirkungsorientierung. Zum einen soll die Umsetzung neuer Unterrichtsmethoden analysiert werden, um aus positiven und negativen Effekten Schlüsse ziehen zu können. Zum anderen sollen die Lehrer- und Schülerleistungen im ganzen Land systematisch erfasst werden, damit sichtbar wird, ob die erwarteten Wirkungen die Klassenzimmer erreichen. Es

wird vorgeschlagen, dass Praxisschulen als Modelle für die verbesserten Unterrichts- und Lernpraktiken eingesetzt werden.

Die Konzentration eines Vorhabens auf die umfassende Kapazitätsentwicklung von Schulämtern und eine Stärkung ihrer Rolle im Kontext der Schulaufsicht und Beratung ist ein neuer Ansatz, der im Vorhaben in Benin umgesetzt wird. Zu Beginn des Vorhabens wurde eine aufwändige Organisationsdiagnose durchgeführt. Die partizipative Vorbereitung und Durchführung in rund 30 Schulämtern im ganzen Land stellte eine wichtige Etappe für die erfolgreiche Implementierung des Vorhabens dar. Diese Maßnahme wurde allerdings als zu zeitaufwändig und zu teuer kritisiert. Mit dem Konzept eines leistungsfähigen Schulamtes und durchgeführten Schulungen leistet das Vorhaben einen wesentlichen Beitrag zur Verbesserung der Dienstleistungen auf der Ebene der Schulämter. Alle Personalkategorien werden geschult, auch die Sachbearbeiter/-innen. Erste sichtbare Verbesserungen werden z.B. bei Aufnahme/Empfang sowie im Umgang mit Post konstatiert, was zu einer erhöhten Zufriedenheit und deutlich gesteigener Arbeitseffizienz führte. Existierende und anerkannte Instrumente wie z.B. zur Kontrolle der Anwesenheit der Lehrkräfte wurden reaktiviert, andere Instrumente wie z.B. Unterrichtsbeobachtungsbögen um Genderaspekte ergänzt. Damit soll die Durchsetzung der Schulpflicht und die Anwesenheit der Lehrer in streikfreier Zeit erhöht werden. Allerdings wurden neue Instrumente noch nicht ausreichend entwickelt oder Innovationen pilotiert und auf Übertragbarkeit analysiert. Netzwerke zur Förderung des Austauschs wurden für alle 85 Schulamtsleiter in den sechs Departements umgesetzt, über die alle Schulämter erreicht werden. Die Anzahl der beteiligten Personen erweist sich im Falle der Netzwerke für die Gruppe der pädagogischen Berater/-innen aber als zu groß und der Austausch findet nur unsystematisch statt. Als besonders effektiv hat sich das auf die Fortbildungen fußende berufsbegleitende Coaching unter Zuhilfenahme einer Gruppe engagierter pensionierter Schulinspektoren in den Schulämtern herausgestellt.

3.2.2.3 Dezentrale Evaluierungen

Aus den beiden dezentralen Evaluierungen lassen sich nur indirekte Schlussfolgerungen zu sektorspezifischen bzw. fachlichen Erfolgs- und Misserfolgskriterien ableiten. Im Kosovo fanden befragte Jugendliche, dass ihre Bedürfnisse im Bereich formaler Bildung sehr gut und im Bereich non-formaler Bildung gut abgedeckt werden. Hingegen wünschten sie sich breitere Unterstützung und mehr Angebote im Bereich der Berufsberatung und Stellenvermittlung, dem Zugang zu Justiz und Jugendrechten sowie dem Zugang zu Auslandserfahrungen und Jugendaustauschprogrammen.

Im Vorhaben Friedenserziehung in Zentralamerika und Kolumbien wird von individuellen Veränderungen bei Teilnehmenden, Fach- und Führungskräften aus staatlichen und nichtstaatlichen Organisationen des Bildungssektors berichtet. Sie hinterfragten eigene Glaubenssätze und die eigene Realität, bauten Feindbilder ab, erlernten einen respektvollen Blick für das Gegenüber und erarbeiteten alternative Handlungsoptionen. Ihre veränderte Haltung und ihr Wissen geben die ehemaligen Teilnehmenden in ihren Organisationen weiter. Die Verbreitung von Instrumenten zur Konfliktlösung sei dabei über die institutionelle Ebene hinaus auch auf regionaler Ebene wahrzunehmen.

3.2.3 Querschnittsthemen

3.2.3.1 Unabhängige Evaluierungen

Das Querschnittsthema Konfliktbearbeitung / Konfliktprävention findet in den unabhängig evaluierten Vorhaben keinen Eingang. Eine Ausnahme bildet das SV „Innovative Ansätze“ in dem in Zusammenarbeit mit dem SV „Bildung und Konfliktbearbeitung“ zwei konfliktbezogene Studien erarbeitet wurden. Demgegenüber wird der Geschlechtergleichberechtigung in allen UEs Rechnung getragen. Für die Vorhaben südlich der Sahara, bei denen es sich in allen Fällen um Vorhaben des ehemaligen Unternehmensteils GTZ handelt, lässt sich feststellen, dass die Thematisierung des Genderaspekts, trotz noch vorhandener Probleme, als Erfolg bezeichnet werden kann. Dies gilt besonders in Bezug auf eine erhöhte Einschulrate sowie einer geringeren Schulabbruchrate von Mädchen. In den Vorhaben in Ghana und Namibia hat vor allem die konsequente Thematisierung des Genderaspekts in Schulmaterialien zu einer Verhaltensänderung beigetragen. Als besonders gutes Beispiel zum Querschnittsthema Gender ist das Vorhaben in Guinea zu nennen. Es wurde eine Verhaltensänderung bei Mädchen und ihren Eltern bezüglich Bildung festgestellt, Mädchen wurden selbstbewusster und möchten vermehrt im Schulsystem verbleiben. Diese Erfolge scheinen sich besonders auf ein System kleiner Klassen, den Einbezug insbesondere der Mütter, einen Fokus auf Mädchenklassen und gut ausgebildete Lehrerinnen zurückführen zu lassen. Die Disparitäten bei der Einschulung von Mädchen konnten zugunsten der Mädchen verringert werden, wie zum Beispiel in Labé, wo der Gender Parity Index auf 0.9 erhöht werden konnte. Auch die Verbleibsrate bei Mädchen in der Primarschulbildung konnte leicht gesteigert werden. Man nimmt an, dass das Vorhaben in Guinea zu diesen Verbesserungen beigetragen hat.

In den durch den Ex-Unternehmensteil InWEnt durchgeführten Vorhaben in Lateinamerika scheint das Querschnittsthema Geschlechtergleichberechtigung durch eine explizite Förderung von Frauen sowie die Einbindung von Genderaspekten in die pädagogische Arbeit ebenfalls positive Wirkungen zu zeigen. In Pakistan verhalf der Genderfokus Mädchen zu einem verbesserten Zugang zu formaler sowie auch nonformaler Bildung.

Im SV „Innovative Ansätze“ war das Thema der Geschlechtergleichberechtigung zuerst Haupt- und dann Nebenziel. Trotzdem wurde das Thema im Sinne einer Mädchenförderung nach Ansicht der meisten von den Evaluatoren befragten Personen nicht ausreichend berücksichtigt. Dennoch konnte aber die Mädchenförderung in der Konzeption der bilateralen Vorhaben verankert werden. Auch hat die Thematik Eingang in zahlreiche vom SV erstellte Dokumente gefunden.

3.2.3.2 Projektfortschrittskontrollen

Im Rahmen der PFKs ist das Querschnittsthema Konfliktbearbeitung / Konfliktprävention in fünf Vorhaben (Guinea, Kongo, Mosambik, Honduras, Sektorvorhaben Bildung) präsent. Geschlechtergleichberechtigung wird in sieben Projekten behandelt. Im SV werden Krisenprävention bzw. Konfliktbearbeitung als Querschnittsthema zwar bearbeitet, sie sind jedoch kein explizites Ziel des SV. Viele Bildungsvorhaben finden in fragilen Kontexten statt (z. B. Guinea – Militärputsch 2008), weshalb die Evaluatoren darauf hinweisen, dass Instrumente, Methoden und Konzepte konfliktensibel entwickelt und umgesetzt werden müssen. Gemäß den Evaluatoren zeigt ein Fokus auf eine verbesserte Bildung zudem gerade in fragilen Kontexten eine friedensfördernde Wirkung. Im Vorhaben in Guinea, wo Konfliktbearbeitung / Konfliktprävention zwar kein explizites Ziel darstellt, wird genau dieser Ansatz verfolgt. Durch die Stärkung von Bildungseinrichtungen sollen friedensfördernde Aktivitäten gestärkt werden und zwar vor allem durch einen aktiven Einbezug der Zivilgesellschaft. In den beiden Vorhaben in Mosambik sowie im

Kongo sind die Förderung der Grundbildung sowie der Berufsbildung integraler Teil der Entwicklungsmaßnahme. Diese Kombination scheint sich positiv auf die Konfliktprävention auszuwirken, da vor allem Berufsbildungsmaßnahmen der Zielgruppe berufliche Perspektiven eröffnen, was langfristig friedensfördernd wirken kann. Im Kongo ist die Reduktion von Konflikten ein explizites Ziel und die Evaluatoren kommen zum Schluss, dass die Maßnahmen des Vorhabens eine entsprechende Wirkung zeigen. In Mosambik wird darauf hingewiesen, dass Beschäftigung und Einkommen – beides durch das Vorhaben gefördert – schlussendlich helfen das soziale Konfliktpotenzial zu reduzieren. Für Honduras wird von den Evaluatoren beanstandet, dass das Vorhaben trotz hohem Niveau an Gewalt und geringen Kapazitäten für die Konfliktbearbeitung im Land kein *Peace and Conflict Assessment* anwendet.

Beim Querschnittsthema Geschlechtergleichberechtigung verfolgen die PFKs ähnliche Ansätze wie die unabhängig evaluierten Vorhaben. Durch die konsequente Thematisierung von Genderaspekten sollen Verhaltensänderungen hinsichtlich der gesellschaftlichen Rolle von Mädchen und Frauen bewirkt werden. In Benin, welches als Hauptziel hat, die Leistungsfähigkeit der Schulämter in verschiedenen Bereichen zu stärken, bezieht sich der Gender-Ansatz auf die institutionelle Ebene, zum einen eine Thematisierung von sexuellen/anderen Übergriffen auf Kinder (insbesondere Mädchen) und die Verbesserung der Schulabbrecher-/innenrate. Da die Schulabbruchsrate von Mädchen besonders hoch ist, profitieren sie im besonderen Maße von dieser Maßnahme. Gemäß den Evaluatoren floss in der jüngeren Vergangenheit ein Großteil der geberfinanzierten Grundbildungsmaßnahmen in die Mädchenförderung. Heute wird jedoch eher der Ansatz des *gender mainstreaming* verfolgt und nur noch sehr wenige Maßnahmen zielen ausschließlich auf Mädchenförderung ab. Aus der Kampagne *Toutes les filles à l'école* wurde die Aktion *Tous les enfants à l'école*.

3.2.3.3 Dezentrale Evaluierungen

Zum Querschnittsthema Konfliktbearbeitung / Konfliktprävention lassen sich nur im Vorhaben in Zentralamerika relevante Aussagen finden. Das Vorhaben setzt auf die Sensibilisierung auf Mitterebene in Form von Bereitstellungen von Konzepten, Methoden und Materialien sowie der Förderung gegenseitigen Erfahrungsaustauschs.

Der Ansatz im Bereich der Geschlechtergleichberechtigung im Kosovo konzentriert sich auf eine Bewusstseinsbildung von genderrelevanten Problemen mit dem Ziel Stereotypen, Vorurteile und Geschlechterungleichheiten zu verringern. Im Vorhaben in Zentralamerika wird die Geschlechtergleichberechtigung durch Aktivitäten in Themengebieten wie Männlichkeit, sexuellem Missbrauch sowie häuslicher Gewalt Rechnung getragen. Zudem werden durch das Vorhaben sowohl männliche wie auch weibliche Change Agents qualifiziert, die gemäß den Evaluatoren in der Lage sind, Geschlechterfragen weitaus sensibler zu behandeln als vor ihrer Teilnahme an der Maßnahme.

3.2.4 Übergreifende Fragen

3.2.4.1 Unabhängige Evaluierungen

Wirkungsnachweise:

Wirkungsnachweise auf Zielgruppenebene werden in den DAC-Kriterien Effektivität (Ebene der direkten Wirkung) und Impact (Ebene der indirekten Wirkung) beschrieben. Die folgende Grafik bietet eine Übersicht, um die Bewertungen von Effektivität und Impact der Evaluierungen übergreifend beurteilen zu können:

UE	GHA		GUI		HGP		INA		NAM		PAK		PER		CHA		ZA		SV		Ø	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A		
Effektivität	2	ZM	2	ZM	2	ANG	2	ZM	4	ANG	2	ANG	2	ZM	3	ANG	3	ANG	2	ZM	2.4	ANG 5, ZM 5
Impact	3	ZM	2	ZM	3	ANG	3	ZM	4	ANG	3	ANG	3	ZM	3	ANG	3	ZM	2	ANG	2.9	ANG 5, ZM 5

Tabelle 8: Übersicht über die Bewertungen der DAC-Kriterien Effektivität und Impact für UEs

Die durchschnittliche Benotung der DAC-Kriterien Effektivität und Impact beläuft sich auf 2.4 für die Effektivität, respektive 3 für den Impact. Gemäß der GIZ-Benotungsskala stufen die Evaluatoren alle Vorhaben bis auf Namibia als erfolgreich (Noten 1-3) ein, was bedeutet, dass die Vorhaben sowohl auf direkter als auch auf indirekter Ebene Wirkung erzielen sollten. Im Rahmen der Metaevaluierung untersuchten wir die Benotungen auf deren Angemessenheit und kamen zum Schluss, dass jeweils in fünf Fällen die Bewertungen zu mild ausgefallen sind.

Die Gründe für die in vielen Fällen zu milde benotete Bewertung anhand der GIZ-Richtlinien liegen aus unserer Sicht in erster Linie an den nur in seltenen Fällen nachweisbaren Wirkungen auf Zielgruppenebene. Ist das Projektziel wie beispielsweise in Ghana die Verbesserung der Lehrerausbildung, handelt es sich bei der definierten Zielgruppe jedoch um Primarschüler, so ist das Projekt design nicht darauf ausgelegt, direkte Wirkung bei der Zielgruppe zu erreichen. Ein anderes Beispiel hierfür ist das Vorhaben in Peru, wo auf der Ebene der direkten Wirkung explizit keine zielgruppenrelevante Wirkung angestrebt wird, jedoch auf Mittler-Ebene direkte Wirkung in Form von Professionalisierung von Lehrerausbildnern und Lehrern erzielt werden kann. Das SV „Innovative Ansätze“ wurde auf Mittlerebene angesiedelt und Wirkungen bei der Zielgruppe sollten daher über Evaluierungen von bilateralen Vorhaben bewertet werden. Ein weiterer Grund für eine fehlende belegbare Wirkung auf direkter Ebene sind oftmals fehlende Baselinedaten und Vergleichsgruppen, welche einen Wirkungsnachweis des Vorhabens noch erschweren. Eine Übersicht über intendierte und erzielte Wirkung ist im Anhang auf Seite 81 zu finden.

Auf der Ebene der indirekten Wirkung, das heißt unter anderem in den Bereichen Armutsreduzierung, Beschäftigungs-/ Einkommenschancen und Verbesserung der Schülerleistungen, wurde auf Zielgruppenebene in fünf Vorhaben keine indirekte Wirkung erzielt. Die Gründe hierfür sind unterschiedlich. In den Vorhaben in Zentralamerika, Honduras, Guatemala und Peru wurde auf die mangelnde Zeit, einen Impact generieren zu können beziehungsweise einen noch nicht abgeschlossenen Prozess hingewiesen. In Namibia wurde die mangelnde Breitenwirksamkeit des Vorhabens unterstrichen. Für die Vorhaben in Indonesien und in Ghana kamen wir im Rahmen der Metaevaluierung zum Schluss, dass in der Wirkungskette die Impactebene nicht richtig berücksichtigt wurde. Somit können für diese beiden Vorhaben keine Aussagen bezüglich nachweisbarer indirekter Wirkung getroffen werden.

Bildungsqualität:

Bildungsqualität wird durchweg mit guten Schulleistungen der Schüler in Verbindung gebracht. Auf eine verbesserte Bildungsqualität wird geschlossen, wenn Schüler im Schulsystem des jeweiligen Landes Noten erreichen, die ihnen den Übertritt in die nächste Klasse oder von der Primar- in die Sekundarschule erlauben. Auch der Verbleib im Schulsystem, d.h. niedrigere Schulabbruchraten und der Anteil erfolgreicher Primarschulabschlüsse zählen als Faktoren. Zudem werden bessere Noten in bestimmten Fächern als ein Indiz für eine verbesserte Bildungsqualität gewertet. In Ghana stehen z.B. die Schulleistungen in Sprache, Mathematik und Life Skills im Fokus, in Guinea die Leistungen in Französisch und Mathematik.

Zudem wird in einigen Vorhaben aus Voraussetzungen und Leistungen von Lehrern auf die Bildungsqualität geschlossen, da sich diese Faktoren auf das Lernen und die Schulleistungen der Schüler auswirken. In Pakistan wird dazu Ausbildungsniveau und Weiterbildung von Lehrern und ihre Arbeitsbelastung betrachtet. Lehrerleistungen werden hier anhand der Sprechzeit der Lehrer im Unterricht, dem Anteil von Schülern, die Antwort geben, der Verwendung von Unterrichtsmaterial und dem Schüler-Lehrer-Verhältnis beurteilt.

Auch Faktoren insbesondere auf Seiten von Schülerinnen, welche sie vom Schulbesuch abhalten, werden genannt; z.B. im Tschad frühe Verheiratung, ungewollte Schwangerschaften, eine hohe Arbeitsbelastung im Haushalt und/oder als Arbeitskraft im Kleingewerbe. Mangelnde Sensibilität von männlichen Lehrkräften wird hier ebenfalls als Einflussfaktor für den Schulverbleib von Mädchen betrachtet.

Ergänzend werden die Lehrer-Schüler-Relation oder Infrastruktur (z.B. fehlende Latrinen und direkter Zugang zu Wasser im Tschad) und Schulausstattung als Faktoren einbezogen, die Hinweise auf die Bildungsqualität erlauben.

Die Verbesserung der Bildungsqualität ist in allen Bildungsvorhaben eine zentrale Absicht der Bemühungen. Entsprechend zielen alle in den Bildungsvorhaben verwendeten Instrumente auf eine Optimierung der Bildungsqualität ab. So entwickelte und verbreitete beispielsweise das SV „Innovative Ansätze“ inhaltliche und methodische Ansätze zur Verbesserung der Bildungsqualität vor allem im Bereich Capacity Development, ausgehend von dem Verständnis, dass eine quantitative Ausdehnung von Bildung von der Verbesserung der Bildungsqualität begleitet werden muss. Strategien zur Erhöhung der Bildungsqualität konnten dann auch in den GPE-Partnerländern verankert werden und die Evaluatoren kamen zum Schluss, dass das SV dazu einen wertvollen Beitrag leisten konnte.

Systemrelevanz:

Mit Ausnahme des Vorhabens in Pakistan, das in einer Notfall-Logik operierte, waren alle Vorhaben darauf ausgerichtet auf der Systemebene eine nachhaltige Wirkung zu haben. Doch obwohl den Vorhaben eine hohe Relevanz bezüglich der strategischen Ausrichtung attestiert wurde, konnten sie gemäß den Evaluierungen wenig Wirkung auf Systemebene erzielen. Die in den Evaluierungen festgehaltene fehlende Systemrelevanz spiegelt sich nicht in der positiven Benotung der Nachhaltigkeit gemäß DAC-Kriterium wider, wie untenstehende Tabelle zeigt. Wir sind deshalb der Meinung, dass in einigen Fällen zu milde bewertet wurde.

UE	GHA		GUI		HGP		INA		NAM		PAK		PER		CHA		ZA		SV		∅	Angemessenheit der Beurteilung		
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A						
Nachhaltigkeit	3	ANG	2	ANG	3	ANG	3	ZM	3	ZM	2	ZM	3	ZM	3	ZM	3	ZM	3	ZM	3	ANG	2.8	ANG 4, ZM 6

Tabelle 9: Übersicht über die Bewertungen des DAC-Kriteriums Nachhaltigkeit für UEs

Für die fehlende Systemrelevanz wurden verschiedene Gründe aufgeführt, solche, die im Einflussbereich des Vorhabens selbst liegen und solche, die vom Vorhaben wenig beeinflusst werden konnten. Zu ersterem gehört, dass im Tschad die institutionelle Verankerung nicht von Anfang an mitgedacht wurde oder dass in Namibia bei der Unterstützung der Dezentralisierungsprozesse die Stärkung der dezentralen fiskalen Strukturen nicht miteinbezogen wurde. Auch eine zu kurze Laufzeit wurde immer wieder als Begründung für fehlende systemische Nachhaltigkeit aufgeführt. So wurde z.B. in Peru festgehalten, dass für die Unterstützung von komplexen Veränderungsprozessen eine längere Präsenz notwendig gewesen wäre. In Indonesien wurde konstatiert, dass die Projektintervention zu eng und zu „supply-driven“ angelegt gewesen sei, um die angestrebte Breitenwirkung auf die Unterrichtsmethoden der Lehrpersonen zu erzielen. Dazu kam, dass die „Science Kits“ für Lehrpersonen von einem Anbieter in Deutschland hergestellt wurden, was die Beschaffung von neuen Kits oder Ersatzteilen wesentlich erschwerte. Das Vorhaben in Tschad wurde als zu isoliert und autonom bewertet. Für eine Breitenwirkung habe es zu wenig sichtbare Erfolgsbeispiele generieren und verbreiten können.

Das Capacity Building auf individueller Ebene wurde in der Regel als erfolgreich eingestuft. Der geplante Transfer auf die Systemebene hat dann aber oft nicht wie geplant stattgefunden. So gab es in Indonesien nur in vereinzelten Fällen einen Transfer von den vom Projekt ausgebildeten Lehrpersonen zu neuen Lehrpersonen. Die dafür eingerichteten Lehrer-Arbeitsgruppen existierten zwar noch, behandelten aber andere Themen. Auch in Ghana wurde das Capacity Building von Einzelpersonen als gelungen beschrieben. Für eine Wirkung auf Systemebene seien jedoch zu wenige Individuen ausgebildet worden. Dazu kam, dass die jungen frisch ausgebildeten Lehrer und Lehrerinnen wenig Einfluss in den Schulen ausüben konnten. Die älteren, erfahrenen Lehrkräfte schienen wenig offen für Anregungen von jungen und unerfahrenen Lehrpersonen, speziell in einer Gesellschaft, in der das Alter hohes Ansehen genießt. In Namibia fehlte nebst dem individuellen Capacity Building ein systemischer Beratungsansatz. In Zentralamerika hatten die in Deutschland ausgebildeten Fachkräfte wenig Einfluss auf die Führungskräfte in ihren Organisationen. Sie führten zwar teilweise Transferprojekte durch, konnten damit aber keine institutionelle Verankerung von neuen Ansätzen und Themen bewirken. Dazu kam, dass im Projekt Honduras, Guatemala, Peru die Fachkräfte nicht vertraglich verpflichtet wurden, nach ihrer Ausbildung in ihren Organisationen zu verbleiben.

Der häufigste Grund außerhalb des Einflussbereichs der Vorhaben war, dass Reformprozesse zwar auf dem Papier offiziell von Regierungsseite unterstützt wurden, es aber mit der Umsetzung nicht voranging, d.h. dass Absichtserklärungen und Umsetzung nicht übereinstimmten. Es kam auch immer wieder vor, dass von Vorhaben unterstützte und beeinflusste „Policy Papers“ nur sehr verzögert oder gar nicht offiziell verabschiedet wurden. Auch kam es zu Rückfällen in tradierte Praktiken, wie in Honduras, Guatemala und Peru. Bei der Umstellung von einem autoritätsbasierten auf einen informationsbasierten Ansatz fürchteten in Ghana die Regierungspartner um ihren Einfluss.

Die Vorhaben fanden in politisch fragilen Kontexten statt. In Peru wurden Dezentralisierungsbestrebungen nach einem Regierungswechsel rückgängig gemacht, so dass die strategischen Vorstellungen des Vorhabens und der Regierungen auseinanderdrifteten. In Guatemala und Honduras war nach dem Regierungswechsel das Interesse am Vorhaben seitens der Regierung nicht mehr gegeben und das Vorhaben war in den zuständigen Ministerien wegen häufigen Personalwechsels kaum mehr bekannt. Allerdings kann sich ein Regierungswechsel auch positiv auswirken, so geschehen bei einem erneuten Regierungswechsel in Peru.

Nebst den aufgeführten kritischen Punkten gab es in den UEs aber auch hoffnungsvollere Kommentare. So war in Guinea durch den Einbezug der Zivilgesellschaft und der Generierung von Ownership für die dezentralen Bildungspläne eine nachhaltige Stärkung des dezentralen Bildungssystems gelungen. Zudem konnten durch Restrukturierungen, die vom Vorhaben unterstützt wurden, Kompetenzstreitigkeiten zwischen Abteilungen in der Regierung reduziert werden. In Peru und Guatemala waren die Partneruniversitäten dank ihrer Experimentierfreudigkeit offen für die Umsetzung von Innovationen. In Ghana trug das Vorhaben zu einer nachhaltigen Sprachenpolitik bei. In Namibia hinterließ das Vorhaben gültige und genutzte Richtlinien für Schulinspektoren und Elternbeiräte sowie ein funktionierendes Schulverbundsystem, das jedoch wegen mangelnder Unterstützung und fehlender Richtlinien gefährdet war. Im Tschad entstand eine gewisse Nachhaltigkeit über ehemalige Projekt-Mitarbeitende, die nach Vorhabenende ihre im Projekt erworbenen Kenntnisse und Erfahrungen in andere Organisationen einbrachten. In Ghana besetzten drei ehemalige Counterparts nach Vorhabenende weiterhin wichtige Funktionen mit Entscheidungskompetenz in der Partnerstruktur.

Positive Aspekte waren auch für das SV „Innovative Ansätze“ zu vermerken. Das Sektorvorhaben habe dazu beigetragen, dass Deutschland international als anerkannter und fachlich kompetenter Partner in der Bildungszusammenarbeit wahrgenommen wurde. Dies konnte einerseits durch die Erarbeitung und Verbreitung der Capacity Development Guidelines - die auch noch fünf Jahre danach Verwendung fanden - und andererseits über die Verankerung der Themen Bildungsqualität, frühkindliche Bildung und Mädchenförderung bei den zentralen Partnerorganisationen GPE und ADEA gelingen. Trotzdem konnte eine nachhaltige Vorreiterrolle Deutschlands aufgrund der unzureichenden Priorisierung der Grundbildung (abnehmende Fördermittel) seitens des BMZ und der damit verbundenen eingeschränkten Umsetzungsmöglichkeiten nicht konstatiert werden. Ein weiterer Misserfolgswert für dieses SV war das hohe Maß an kurzfristigen Beratungsaufgaben für das BMZ und der damit zusammenhängenden Vernachlässigung von langfristigen konzeptionellen Aufgaben.

Leistungsmix:

Zum Leistungsmix lässt sich UE-übergreifend feststellen, dass sich die Vorhaben vergleichbarer Ansätze bedienen, innerhalb der Vorhaben werden jedoch Leistungen unterschiedlich kombiniert. So wurden in Ghana beispielsweise Fortbildungen für Lehrpersonen sowie ausgewählte Stakeholder mit der Entwicklung von Materialien und Methoden eines Teacher-Training-Konzepts und Maßnahmen zur Organisationsentwicklung und Netzwerkbildung kombiniert. In Guinea wurden Qualifizierungsmaßnahmen für Lehrpersonen und anderes Schulpersonal mit diversen Beratungsleistungen sowie der Erarbeitung eines Lehrerausbildungskonzepts bzw. -profils für FIERE-Lehrerinnen kombiniert. In Namibia wurden u.a. Fortbildungen von Lehrpersonen zur Anwendung eines neuen

Curriculums mit unterstützenden Beratungen zur Organisationsentwicklung bei den Bildungsministerien und Beratungsleistungen zur Einführung eines neuen Curriculums verbunden.⁸ In Zentralamerika wurden u.a. Fort- und Weiterbildungsmaßnahmen von Fach- und Führungskräften mit der Beratung von Mittlerorganisationen zum Ausbau der institutionellen Fähigkeiten zur friedlichen Konfliktlösung sowie der Erstellung eines Kurskonzepts für E-Learning-Komponenten im Bereich Friedenspädagogik kombiniert. Im Tschad wurden unter anderem Schulungen für Lehrpersonen und andere Personen des Bildungskontextes mit Organisationsberatungen von Elternvereinen (insbesondere in Fragen der Verwaltung) und Beratungen zur Vorbereitung eines Bildungsprogrammes verbunden.

Die Erbringung der einzelnen Leistungen wird meist als erfolgreich bewertet. Dabei zeigte sich, dass auch sehr vielversprechende Ansätze und Leistungen häufig erst im Zusammenspiel mit anderen Maßnahmen ihre Wirkung entfalten konnten, gleichzeitig ist der wirkungsvolle Mix von Leistungen ein sehr herausforderndes Unterfangen. Für das Engagement der Partner erwies es sich dabei als förderlich, wenn es dem Vorhaben gelang, an Bestehendes anzuknüpfen, wie beispielsweise im Tschad an die Tradition von Elternvereinen. Die Datengrundlage lässt jedoch kaum Schlüsse in Bezug auf universell wirksame Leistungsmixe zu, da die Leistungsmixe einerseits stets an den Kontext angepasst sind und andererseits in ihrer Ausgestaltung variieren. Am ehesten als übergreifend erfolversprechend lässt sich die Verbindung von Lehrerfortbildungen mit der Entwicklung und Einführung neuer Lehrmittel sowie die Verwendung eines Mehrebenenansatzes identifizieren.

Geberkoordination:

Hinsichtlich der Geberkoordination lässt sich feststellen, dass eine Koordination mit anderen bilateralen sowie mit multilateralen Gebern mit Ausnahme der Vorhaben in Honduras, Guatemala und Peru in unterschiedlichem Ausmaß in allen ausgewerteten UEs stattfindet. Als ein zentraler Punkt wird eine verstärkte Abstimmung der Geber mit den Bildungsministerien angesehen, da es sonst entweder zu parallelen Strukturen beziehungsweise Überschneidungen von Leistungen oder beschränktem Einfluss der Geber im politischen Reformprozess kommen kann. Die Zusammenarbeit mit anderen Gebern wird in den Vorhaben Ghana, Guinea, Indonesien und Namibia grundsätzlich als positiv oder eher positiv bewertet. Verbesserungspotenzial bezüglich Überschneidungen und mangelndem Einfluss der Geber auf politischer Ebene findet sich in den Vorhaben in Pakistan, Tschad und Zentralamerika.

Rückschlüsse auf die Kooperation von finanzieller Zusammenarbeit (FZ) und technischer Zusammenarbeit (TZ) lassen sich aus sieben UEs ziehen. Keine direkte TZ/FZ-Zusammenarbeit bestand in Peru sowie in Pakistan. Vorhabenübergreifend lässt sich auf eine qualitativ gute Zusammenarbeit zwischen TZ und FZ schließen. Explizit positiv beurteilt wird die Zusammenarbeit in den Vorhaben in Ghana, Guinea und im Tschad, wo von einer perfekten Zusammenarbeit aus einem Guss gesprochen wird. Eine sehr enge Zusammenarbeit bestand auch beim Vorhaben in Indonesien, wobei es sich um ein Gemeinschaftsprojekt von TZ und FZ handelte. Interessant in diesem Fall ist, dass sowohl TZ als auch FZ das Vorhaben separat evaluierten und besonders in Bezug auf Effizienz zu sehr unterschiedlichen Ergebnissen gekommen sind. Eine Analyse gemäß dem DAC-Kriterium Effizienz ergab in der Ex-post Evaluierung der KfW den Wert 4, die Schlussevaluierung

⁸ Es muss an dieser Stelle einschränkend erwähnt werden, dass die einzelnen Maßnahmen – insbesondere in den PFKs – nicht immer detailliert in den Evaluierungen ausgeführt wurden. So war es für uns nicht immer klar ersichtlich, welche Leistungen tatsächlich miteinander kombiniert wurden.

der TZ bewertete die Effizienz mit der Note 1 – sehr gut. Bezüglich der TZ/FZ-Zusammenarbeit wird das Vorhaben in Namibia zwar als positiv bewertet, doch auch in diesem Fall wird beanstandet, dass die Kooperation zwischen TZ und FZ Effizienzmängel aufweist.

3.2.4.2 Projektfortschrittskontrollen

Wirkungsnachweise:

In einem ersten Schritt werden wir auch zu den PFKs die relevanten Feststellungen bezüglich der Wirkungsnachweise erläutern. In der untenstehenden Tabelle wird ersichtlich, dass sich die durchschnittlichen Benotungen der DAC-Kriterien Effektivität und Impact anhand der GIZ-Richtlinien auf 2,6 für die Effektivität und 2,9 für den Impact belaufen. Diese Werte sind praktisch identisch mit den Benotungen der unabhängig evaluierten Vorhaben. Demzufolge werden die PFKs für die Kriterien Effektivität und Impact PFK-übergreifend ebenfalls als erfolgreich betrachtet.

PFK	BEN		GUI		HON		COD		MAW		MTN		MOZ		SV		ZAs		Ø	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A				
Effektivität	3	ZM	2	ZM	2	ZM	2	ZM	3	ANG	2	ZM	3	ANG	3	ANG	2	ANG	2.4	ANG 4, ZM 5
Impact	2	ZM	3	ZM	3	ZM	3	ZM	3	ANG	3	ZM	4	ANG	3	ANG	2	ANG	2.9	ANG 4, ZM 5

Tabelle 10: Übersicht über die Bewertungen der DAC-Kriterien Effektivität und Impact für PFKs

Ähnlich wie bei den UEs bewerteten wir die Angemessenheit der Benotungen der Evaluatoren. Bei den PFKs sind wir der Ansicht, dass von sechzehn Benotungen anhand der DAC-Kriterien Effektivität und Impact acht zu mild bewertet wurden, da auf Zielgruppenebene nur selten eine belegbare Wirkung nachgewiesen werden kann.

In den Tabellen zur Wirkungsmessung auf Zielgruppen- und Mittlerebene auf Seite 81 werden die erzielten Wirkungen der PFKs dargestellt. Auf der Ebene der direkten Wirkung gelingt es den PFKs mehrheitlich (in sechs von acht Vorhaben) eine Wirkung für die Zielgruppen nachvollziehbar darzustellen. Bis auf das Vorhaben im Kongo, wo die direkte Wirkung auf Zielgruppenebene quantitativ nachweisbar ist, basieren die Ergebnisse aus den Vorhaben Benin, Guinea, Malawi, Mauretanien und Zentralasien auf qualitativen Aussagen der Evaluatoren. Bezüglich der indirekten Wirkung sieht das Resultat etwas weniger positiv aus. Nur in Guinea lässt sich anhand der Beschreibungen auf Impact Ebene für die Zielgruppe eine Wirkung feststellen, ansatzweise ist dies auch im Kongo noch der Fall. Dieses Resultat ist jedoch verständlich, da es sich bei den vorliegenden Projekten noch um laufende Vorhaben handelt. Des Weiteren wurde von den Evaluatoren der PFKs auch mehrmals darauf hingewiesen, dass es noch zu früh sei, eine Wirkung auf der Impact Ebene zu lokalisieren.

Bildungsqualität:

Auch bei den PFKs wird durchweg aus den Leistungen der Schüler auf Bildungsqualität geschlossen. In Malawi werden z.B. die Lernleistungen der Schüler anhand der Schulnoten in Rechnen, Lesen und Schreiben beurteilt. In Mosambik werden dazu die Leseleistungen der Schüler gemessen. Zudem werden dieselben Faktoren wie bei den UEs als Hinweise auf Bildungsqualität genannt.

Auch bei allen PFKs ist die Verbesserung der Bildungsqualität in allen Bildungsvorhaben eine zentrale Absicht der Bemühungen. Entsprechend zielen alle in den Bildungsvorhaben verwendeten Instrumente auf eine Optimierung der Bildungsqualität ab.

Systemrelevanz:

Auch hier sind die evaluierten Vorhaben auf Systemwirkung ausgerichtet. Die Ausnahmen bilden das Vorhaben im Kongo, das aufgrund des schwierigen Umfeldes einer Notfall-Logik folgt sowie das Vorhaben Sektorberatung, das auf einer anderen Ebene anzusiedeln ist.

Die Kommentare in den PFKs zeigen zwar auch hier ein kritisches, jedoch im Vergleich zu den UEs etwas hoffnungsvolleres Bild. Wahrscheinlich gerade weil eine nachhaltige Systemrelevanz in einem laufenden Vorhaben noch wenig abschätzbar ist, hoffte man durch geeignete entwicklungs- politische Ansätze und Maßnahmen und durch Teilerfolge ermutigt, die angestrebten Wirkungen zu erreichen. Die Kommentare basieren zum Teil auch auf einer Einschätzung von Wahrscheinlichkeiten und Potenzialen. Interessanterweise ist dieses im Vergleich zu den UE etwas positivere Bild nicht in den Benotungen der Nachhaltigkeit abgebildet (siehe untenstehende Tabelle). Wir schätzen diese Benotungen darum als angemessener ein als diejenigen in den UEs.

PFK	BEN		GUI		HON		COD		MAW		MTN		MOZ		SV		ZAs		∅	Angemessenheit der Beurteilung
	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A		
Nachhaltigkeit	3	ANG	3	ANG	3	ZM	4	ANG	3	ANG	3	ZM	4	ANG	3	ANG	2	ZM	3.1	ANG 6, ZM 3

Tabelle 11: Übersicht über die Bewertungen des DAC-Kriteriums Nachhaltigkeit für PFKs

In Zentralasien wurde konstatiert, dass eine Übernahme von Innovationen durch das Bildungsministerium teilweise stattgefunden habe, diese jedoch noch nicht landesweit eingesetzt wurden. Die Perspektiven für den regionalen Austausch wurden als positiv eingeschätzt. Auch wurde gewürdigt, dass nebst dem individuellen auch ein institutionelles Capacity Development gemacht wurde. In Benin hat der Bildungssektor eine Pilotfunktion für den Dezentralisierungsprozess, das Vorhaben berät die de-konzentrierte Bildungsverwaltung im Zuge dieses Dezentralisierungskurses und unterstützt deren Beratungsrollen gegenüber den Kommunen. In Mauretanien gelang es alle gewählten und angestellten Kommunalvertreter/-innen zu schulen und auf diese Weise den Dezentralisierungskurs zu unterstützen. Das Vorhaben in Malawi hatte Büros im Ministerium und dadurch Zugang zu wichtigen Entscheidungsträgern. Im Ministerium in Mosambik konnte eine Schulentwicklungsplanung, ein Planungs-, Monitoring und Evaluierungszyklus eingeführt und Trainingsmodule formalisiert werden mit guten Aussichten auf Nachhaltigkeit. Es wurde auch betont, dass eine nachhaltige Wirkung nicht im Rhythmus kurzer Programmphasen zu erreichen, dass diese vielmehr das Ergebnis einer langjährigen intensiven Zusammenarbeit sei. Als sehr erfolgreich wurde das Vorhaben in Guinea eingestuft. Dort wurden die vom Vorhaben initiierten Prozesse auf nationaler Ebene formalisiert. Dem Vorhaben sei es gelungen, eine Verbindung von der Basis zur strukturellen Ebene herzustellen.

Wie bereits erwähnt, gab es auch in den PFKs einige kritische Anmerkungen bezüglich der Systemrelevanz. Ein immer wiederkehrendes Thema war der fragile, politische Kontext, der eine Wirkung auf Systemebene immer wieder zu Nichte machte und von den Vorhaben nur schwerlich beeinflusst werden konnte. Beispiele dafür sind Honduras, Mauretanien, Zentralasien oder der Kongo. So blieb der Bildungssektor in Honduras schwach trotz jahrelanger EZ-Unterstützung und die

Regierung hatte keine klare Dezentralisierungsstrategie. Es waren kaum Bemühungen der Regierung sichtbar, regulative Rahmenbedingungen zu genehmigen und die Anwendung einer nationalen Strategie blieb begrenzt. Auch in Mauretanien wurden die formal vollzogenen Reformen nicht mit Leben gefüllt. Dies ist teilweise zurückzuführen auf einen Konflikt zwischen den Prinzipien der Dezentralisierung und der Kontrolllogik des Staates. Gemäß den Evaluatoren wurde ein gemeinsamer Ansatz aus der Taufe gehoben, als die Rahmenbedingungen für Reformprozesse noch völlig andere waren. Auch in Malawi bestand die größte Herausforderung darin, dass formale Strategien gar nicht oder in einer anderen Geschwindigkeit als vom Projekt erwartet umgesetzt wurden. Dies wurde mit den ungenügenden Umsetzungskapazitäten im Ministerium, einer schwachen Motivation der Beamten, vieler Absenzen durch Teilnahme an Workshops, häufigen Personalwechseln und längeren Vakanzen in wichtigen Positionen begründet. Es sei zudem nicht gelungen, die Partner vom Wert einer Planung zu überzeugen. In Zentralasien war die fragile politische Konstellation ein Unberechenbarkeitsfaktor. Für Mosambik wurde festgestellt, dass die Wirkungsketten lang seien und dass sich Arbeitsweisen und Verhalten von Lehrkräften - wie überall - nur sehr langsam verändern würden. Aktuell könne daher noch nicht von Impact gesprochen werden. In Benin sind viele der Schulumtsleitungen nahe des Pensionsalters, so dass erworbenes Wissen nicht mehr lange im System bleiben wird.

Leistungsmix:

Zur Frage nach dem Leistungsmix im Rahmen der PFKs stellen wir fest, dass wie bei den unabhängigen Evaluierungen die Kombination aus verschiedenen Leistungen von Projekt zu Projekt je nach Auftrag und Problemstellung divergiert. Ähnlich wie bei den UEs scheint auch bei den PFKs die Schwierigkeit weniger in der Erbringung der einzelnen Leistungen zu liegen, sondern ist vielmehr in deren Kombination zu lokalisieren.

Geberkoordination:

Eine Geberkoordination mit anderen bilateralen sowie mit multilateralen Gebern fand in unterschiedlichem Ausmaß in allen untersuchten Vorhaben statt. Wie im Fall der UEs fordern auch die Evaluatoren von PFK-Vorhaben eine verstärkte Geberabstimmung, da es sonst zu parallelen Strukturen beziehungsweise Überschneidungen von Leistungen kommen kann. Überschneidungen wurden im Vorhaben in Honduras generell bemängelt und im Vorhaben in Benin spezifisch im Bereich der Fortbildungsangebote. Des Weiteren ist in Malawi auf Partnerseite eine Skepsis gegenüber der Geberkoordination festgestellt worden. Im Vorhaben in Zentralasien befindet sich die Geberkoordination noch in den Anfängen, soll aber in den unterschiedlichen Projektländern weiterentwickelt und ausgebaut werden. In Mauretanien wird die Rolle des verantwortlichen Ministeriums kritisiert, welches eigentlich die Aufgabe hätte die Geberkoordination zu leiten, diese Aufgabe aber ungenügend wahrnimmt. In Mosambik war der deutsche Beitrag zu Beginn des Vorhabens innerhalb des von der Weltbank finanzierten Projekts angesiedelt, dies wurde jedoch im Laufe des Vorhabens zu einem Hemmfaktor, weshalb ein Wechsel der Durchführungsorganisation vollzogen wurde.

In acht Vorhaben besteht eine Kooperation zwischen TZ und FZ. In Mauretanien hat sich die FZ aus dem Vorhaben zurückgezogen. Vorhabenübergreifend lässt sich auf eine qualitativ gute Zusammenarbeit zwischen TZ und FZ schließen. Ausmaß und Art der TZ/FZ-Zusammenarbeit sind

von Vorhaben zu Vorhaben unterschiedlich. Von einer exzellenten Zusammenarbeit wird besonders im Vorhaben in Guinea gesprochen, wo Partner der Deutschen EZ selbst in der Zeit von politisch sehr schwierigen Verhältnissen eine effiziente Kooperation mit den Partnern vor Ort bescheinigt haben. Schwierigkeiten bei der TZ/FZ-Zusammenarbeit scheinen unter anderem in Malawi zu bestehen, wo sich die KfW nicht an der Evaluierungsmission des Vorhabens beteiligen wollte, was von den Evaluatoren sehr bedauert wurde.

3.2.4.3 Dezentrale Evaluierungen

Wirkungsnachweise:

Hinsichtlich der Wirkungsnachweise der beiden dezentral evaluierten Vorhaben wurde seitens der Evaluatoren keine Benotung der DAC-Kriterien vorgenommen. Aussagen bezüglich der Benotung und der tatsächlich erzielten Wirkung auf Zielgruppenebene können demzufolge nicht getroffen werden.

Was die erzielte Wirkung auf direkter bzw. indirekter Ebene betrifft, zeigt sich bei den beiden dezentralen Evaluierungen ein etwas besseres Resultat als bei UEs und PFKs (siehe Kapitel 6.9 auf Seite 81). Es lassen sich in den Vorhaben im Kosovo und in Zentralamerika Wirkungen auf Zielgruppenebene in drei von vier Fällen nachweisen. Im Kosovo scheint eine Wirkung für die Zielgruppen auf der indirekten Ebene einzutreffen. Auf direkter Ebene ist keine zielgruppenrelevante Wirkung nachweisbar. Im Vorhaben in Zentralamerika ist bei den Projektabsolventen eine gewisse Wirkung nachweisbar. Inwiefern diese breitenwirksam ist, kann jedoch nicht festgestellt werden.

Bildungsqualität:

Der friedenspädagogische Ansatz in Zentralamerika zielte auf eine grundsätzliche Veränderung der Einstellungen und Haltungen der Zielgruppen im Umgang mit Konflikten ab. Direkte Aussagen zum Verständnis von Bildungsqualität werden nicht gemacht, indirekt lässt sich aber ableiten, dass Bildungsqualität auch am Bildungserfolg beurteilt wird. Bildungserfolg wird jedoch nicht mit Noten und Abschlüssen gemessen sondern anhand sichtbarer Veränderungen der Verhaltensweisen der Zielgruppe beurteilt. Das Augenmerk wird also über Kompetenz hinaus auf Performanz gelegt.

Im Bildungsvorhaben wurden wiederum alle Instrumente mit der Zielsetzung der Verbesserung der Bildungsqualität im Sinne dieses Projekts eingesetzt.

Systemrelevanz:

In Zentralamerika wurde bezüglich Reformprozessen seitens der Regierungspartner eine Diskrepanz zwischen Ankündigung und Umsetzung festgestellt. Häufig wechselnde Entscheidungsträger sowie eine fehlende Offenheit und mangelndes Wissen bezüglich Friedenserziehung auf der Regierungsseite erschwerten eine systemische Wirkung des Vorhabens zusätzlich. Das Jugendvorhaben im Kosovo war nicht im Einklang mit den Prioritäten der Regierung.

Das Programm in Zentralamerika war subsidiär zu den Eigenanstrengungen der zivilgesellschaftlichen Partnerorganisationen, jedoch nicht zu denen der staatlichen Partner. Es wurde positiv herausgehoben, dass sich die meisten Teilnehmenden der Kurse für Friedenserziehung als engagierte Multiplikatoren erwiesen. Die Wirkung habe sich von diesen Individuen zuerst auf das persönli-

che, dann auf das Arbeitsumfeld und schließlich auf ein Mikrosystem übertragen. Durch den Austausch von Ehemaligen war ein Friedensnetzwerk mit intensiven Freundschaften entstanden, dessen Fortführung allerdings gefährdet schien. Positiv ist auch zu erwähnen, dass die Partneruniversität in Kolumbien die Kurse zur Friedenserziehung in ihr Lehrangebot aufnahm.

Geberkoordination:

In beiden dezentral evaluierten Vorhaben existieren Geberkooperationen zwischen den Beiträgen aus Deutschland sowie anderen bilateralen und multilateralen Gebern. Im Kosovo wird darauf hingewiesen, dass die Geberkoordination zwischen dem verantwortlichen Ministerium und den Geberinstitutionen zu Beginn des Vorhabens gut funktionierte. Schwierigkeiten tauchten jedoch bei der Umsetzung sowie der Formulierung neuer Projekte auf. Besonders herausgestrichen wird der Mangel an effizienter Geberkoordination bei der Umsetzung der Kosovo Youth Strategie and Action Plan. Im Vorhaben in Zentralamerika lassen sich hinsichtlich der Geberkoordination wenige Informationen ableiten. Es wird jedoch erwähnt, dass das Vorhaben eine gute Ergänzung zu den durch internationale Geber unterstützten Prozessen in der Region darstellt.

In den beiden dezentralen Evaluierungen scheint die TZ/FZ-Kooperation eine geringe Rolle zu spielen. In Zentralamerika wird die gute Zusammenarbeit sowie die Nutzung von Schnittstellen mit dem KfW-Programm CONVIVIR hervorgehoben. Im Kosovo besteht keine explizite Kooperation zwischen TZ und FZ.

Indikatoren für die aggregierte Wirkungsberichterstattung:

Genauere Zahlen hinsichtlich der Anzahl Schüler/-innen und Multiplikatoren, die im Sektor Bildung pro Jahr qualifiziert wurden, können den Vorhaben evaluierungstypübergreifend nicht entnommen werden. Zwar geben vereinzelte Vorhaben wie beispielsweise die UE Indonesien Zahlen der erreichten Schüler für die gesamte Projektdauer an, jährliche Zahlen sind jedoch nicht verfügbar.

4 Fazit

Bewertung der OECD-DAC Kriterien für UEs und PFKs im Durchschnitt

UEs	GHA	GUI	HGP	INA	NAM	PAK	PER	CHA	ZA	SV	Ø
Benotung	2	2	3	2	4	2	3	3	3	2	2.6
Gesamtbewertung											

PFKs	BEN	GUI	HON	COD	MAW	MTN	MOZ	SV	ZAs		Ø
Benotung	2	2	2	3	3	3	4	3	2		2.7
Gesamtbewertung											
Benotung Gesamtbewertung UE- und PFK-übergreifend											2.7

Tabelle 12: Gesamtbewertung der Evaluierungen durch die Evaluatoren

Die von uns ausgewerteten UEs wurden von den Evaluatoren in der Gesamtbewertung im Durchschnitt mit 2.6, die von uns ausgewerteten PFKs im Schnitt mit 2.7 bewertet. Der Durchschnitt der Gesamtbewertung für alle von uns ausgewerteten UEs und PFKs zusammen beträgt 2.7 (vgl. Tabelle 12).

Benotung DAC Kriterium PFK- und UE- übergreifend	Ø
Relevanz	1.6
Effektivität	2.4
Impact	2.9
Effizienz	2.7
Nachhaltigkeit	2.9

Tabelle 13: Bewertung der OECD-DAC Kriterien sowie Gesamtbewertung für alle untersuchten UEs und PFKs

Die untersuchten UEs und PFKs wurden von den Evaluatoren im Durchschnitt anhand der OECD-DAC Kriterien in puncto Relevanz mit 1.6, in puncto Effektivität mit 2.4, in puncto Impact mit 2.9, in puncto Effizienz mit 2.7 und in puncto Nachhaltigkeit mit 2,9 bewertet (vgl. Tabelle 13). Die Gesamtbewertung betrug 2.5. Aus unserer Sicht wurde tendenziell etwas zu milde bewertet.

Manageriale Erfolgs- und Misserfolgskriterien:

Die Kommentare in den Evaluierungen ergaben für alle drei Evaluierungstypen ein Bild von sowohl guten wie auch schlechten Beispielen zur Managementpraxis. Aufgrund der Heterogenität der Vorhaben war es nicht immer möglich, verallgemeinerbare Erfolgs- und Misserfolgskriterien abzuleiten. Trotzdem haben sich ein paar Tendenzen über alle Evaluierungstypen hinweg abgezeichnet. So wurde die Relevanz der strategischen Ausrichtung der Vorhaben positiv beurteilt. Den Vorhaben wurde attestiert, sich mit wichtigen Entwicklungshemmnissen zu befassen und sowohl mit nationalen als auch internationalen Politiken im Einklang zu sein.

Im Bereich der managerialen Erfolgs- und Misserfolgskriterien war die Gestaltung der Kooperationsbeziehungen in den evaluierten Projekten ein wiederkehrendes Thema. Bemängelt wurde die Kooperation hinsichtlich Motivation und Kapazitäten der nationalen Regierungspartner. Die Zusammenarbeit mit Regierungspartnern auf der dezentralen Ebene wurde, wo sie vorkam, allerdings besser bewertet. Ebenfalls positiv beurteilt wurden die Zusammenarbeit mit Nichtregierungspartnern, ihre Motivation und ihre Dynamik, wenngleich diese Punkte ebenfalls nur in einigen Fällen Erwähnung fanden. In einigen Fällen wurden nicht alle relevanten Akteure ausreichend und genügend partizipativ in die Planung einbezogen und/oder deren Motivation und Absichten zu wenig abgeklärt. Wichtig bezüglich der Kooperation scheinen uns folgende zwei Beispiele zu sein: In Guinea funktionierte das Vorhaben auch während der politischen Krise weiter und im Tschad haben Elternvereine eine große Bedeutung und das Engagement und das Interesse der Partner war bezüglich der Unterstützung der Elternvereine dementsprechend sehr hoch. Diese Beispiele zeigen einerseits, wie wichtig Vertrauen und Glaubwürdigkeit bei der Kooperation sind und andererseits, wie es das Engagement der Partner fördert, wenn es dem Vorhaben gelingt an Bestehendes anzuknüpfen.

Für die UEs wurde in einigen Fällen das Programmende thematisiert, das für die nötige Konsolidierung der Ergebnisse oft als zu vorzeitig empfunden wurde. In den PFKs schienen die Kommentare zur strategischen Ausrichtung der Vorhaben etwas positiver auszufallen. Ansonsten konnten wir keine Trendunterschiede zwischen den drei Evaluierungstypen feststellen.

Über alle Evaluierungstypen hinweg tendenziell positiv beurteilt wurde die Relevanz der strategischen Ausrichtung der Vorhaben. Den Projekten wurde attestiert, sich mit wichtigen Entwicklungshemmnissen zu befassen und sowohl mit nationalen als auch internationalen Politiken im Einklang zu sein. Mit Blick auf die Programmenden wurden in den UEs aber oft eine fehlende Exit-Strategie und eine ungenügende Konsolidierung bemängelt.

Sektorspezifische bzw. fachliche Erfolgs- und Misserfolgskriterien:

Im Bereich der sektorspezifischen bzw. fachlichen Erfolgs- und Misserfolgskriterien sind in allen Bildungsvorhaben die Verbesserung des Zugangs und der Bildungsqualität zentrale Anliegen. Vor allem in Subsahara-Afrika stellt der deutlich verbesserte Bildungszugang eine enorme Herausforderung für das Bildungssystem dar, da oft zahlenmäßig ausreichende und qualifizierte Lehrkräfte fehlen. Doch kann eine Verbesserung des Zugangs zu Bildung Erfolge der Anstrengungen zur Verbesserung der Bildungsqualität relativieren und die Effekte dieser Anstrengungen herauszögern. Eine bessere Bildungsqualität wiederum wirkt sich auch positiv auf eine höhere Einschulungs- und geringere Abbruchrate aus. Wenn Bildung als relevant und für die Zukunft der betreffenden Schüler/-innen, ihre Familien und Gemeinden erlebt wird, steigt die Motivation zur Bildungsbeteiligung und beeinflusst den Bildungserfolg positiv. Bildungsqualität und Zugang zu Bildung sind entsprechend Anliegen, die eng miteinander verknüpft sind und weder parallel noch sequenziell, sondern nur systemisch bearbeitbar sind. Dies wurde in allen Bildungsvorhaben unter Berücksichtigung des Kontextes gut berücksichtigt.

In allen Grundbildungsvorhaben werden Maßnahmen zur Stärkung der Lehreraus- und -fortbildung umgesetzt. Es geht dabei um die Verbesserung der fachlichen und der methodisch-didaktischen Kompetenzen zur Erhöhung der pädagogischen Professionalität der Lehrkräfte. Die Entwicklung und Verteilung von neuen Lehr- und Lernmaterialien sowie von Unterrichtsmaterialien sind weitere

wichtige Maßnahmen – im Idealfall wird beides kombiniert. So ist eher gewährleistet, dass Lehrkräfte diese Materialien anwenden können und dies im Unterricht auch tun. Ein gutes Beispiel dafür ist das Projekt in Ghana, bei welchem Lehrmittel in verschiedenen ghanaischen Sprachen unter Einbezug des Personals des Bildungsdepartements sowie der Lehrerbildungsinstitutionen entwickelt wurde. Im Rahmen der Aus- und Fortbildung wird auch häufig eine Verschiebung von lehrerzentriertem zu schülerzentriertem Unterricht thematisiert (z.B. in den Projekten in Ghana, Guinea und Malawi).

Obwohl die Wirksamkeit muttersprachlichen Unterrichts nachgewiesen ist, stößt dieser Ansatz vor allem in einigen afrikanischen Ländern auf Widerstände, weil das Wissen um die positive Wirkung von muttersprachlichem Unterricht auf das Erlernen von Erst- und Zweitsprache sowie auf den Bildungserfolg insgesamt oft fehlt.

Vielversprechend erweist sich die gezielte Verwendung eines Mehrebenenansatzes. Dieser steht dabei in engem Zusammenhang mit der oben genannten Systemrelevanz. Bei einigen UEs wurde deutlich, dass durchaus erfolgreiche Maßnahmen auf einzelnen Ebenen nur geringe oder keine breitenwirksame Wirkung entfalten konnten, weil Faktoren auf anderen Ebenen dies nicht unterstützten. Beispielsweise vermochten verbesserte Unterrichtsmaterialien und die Aus- und Weiterbildung von Lehrkräften in Pakistan keinen schülerzentrierteren Unterricht zu bewirken. Nebst einem zu geringen Vorbildungsniveau der Lehrkräfte behinderten Rahmenbedingungen wie eine schlechte Infrastruktur und zu wenig Unterrichtsmaterial in oft schlechter Druckqualität die Umsetzung eines stärker schülerzentrierten Unterrichts. Andere Beispiele in PFKs deuten darauf hin, dass Maßnahmenbündel, die idealerweise auf Mikro-, Meso- und Makroebene, zumindest aber auf zwei Ebenen ansetzen, effektiver Hebelwirkungen zeigen. Dazu zählt zum Beispiel ein Vorhaben in Zentralasien, in dem in den Lehrerbildungsinstitutionen der Partnerländer Innovationen mit hohem Potenzial für eine nachhaltige Verbesserung der Relevanz der Bildung und der Verwertbarkeit schulischer Lerninhalte eingeführt werden. Politischer Rückhalt für die breitenwirksame Umsetzung wird von Schlüsselakteuren auf der Partnerseite zugesichert und von qualifizierten Institutionen getragen, die durch Beratung und Kapazitätsaufbau durch das Vorhaben in diesem Prozess unterstützt werden.

Eine Reihe evaluierter Projekte operierte in fragilen (Post-)Konfliktsituationen. Bei diesen Beispielen wurden jeweils an die Kontexte angepasste Ansätze verwendet, die der konfliktiven Situation Rechnung trugen, wie beispielsweise die Kombination von Grundbildungsförderung mit non-formaler Berufsbildung in der Demokratischen Republik Kongo oder die Förderung von „Home Schools“ für afghanische Flüchtlinge in Pakistan.

Querschnittsthemen:

Das Thema Geschlechtergleichberechtigung wird in allen drei Evaluierungstypen behandelt. Für die Vorhaben südlich der Sahara, welche durch den Ex-Unternehmensteil GTZ initiiert wurden, lässt sich feststellen, dass die Thematisierung des Genderaspekts als Erfolg bezeichnet werden kann. Dies gilt besonders in Bezug auf eine erhöhte Einschulrate sowie eine geringere Schulabbruchrate von Mädchen. Als besonders wirksam erwiesen sich Ansätze, welche kontextabhängig den spezifischen Bedürfnissen von Mädchen Rechnung trugen und die Zusammenarbeit mit Ältestenräten, religiösen Führern, Elternvertretern sowie den dezentralen Bildungsstrukturen pflegten.

Als gelungenes Beispiel kann dabei ein Projekt in Guinea gelten, in dem die Bildungssituation benachteiligter Bevölkerungsschichten (besonders sozial benachteiligte und lernschwache Mädchen) verbessert werden sollte; HIV & AIDS-Prävention war dabei ein zentrales Querschnittsthema. Das Projekt war dezentral aufgestellt, was sich während einer Krisenzeit, indem eine Zusammenarbeit auf nationaler Ebene nicht möglich war, als Erfolgsfaktor herausstellte.

Zum Querschnittsthema Konfliktbearbeitung / Konfliktprävention lassen sich in fünf PFKs, in einer DE und in einem unabhängig evaluierten Vorhaben relevante Aussagen auffinden. Verallgemeinerbare Aussagen lassen sich aus den sieben Vorhaben, welche dieses Querschnittsthema behandelten kaum ableiten, da die jeweiligen Kontexte und der Ansatz der Durchführungsorganisationen sehr unterschiedlich sind.

Übergreifende Fragen:

Wirkungsnachweise

Im Bereich der übergreifenden, strategisch besonders relevanten Fragen, die seitens der GIZ in den ToR formuliert werden, lässt sich in Bezug auf Wirkungsnachweise zunächst deren oftmaliges Fehlen auf Zielgruppenebene festhalten. Zwar wird in allen Evaluierungen, welche im Rahmen der Synthese beurteilt wurden, plausibel dargelegt, welche direkten und indirekten Wirkungen ein bestimmtes Vorhaben erzielte, jedoch konnten bei weitem nicht alle Evaluierungen Wirkungen auf der Ebene der eigentlichen Zielgruppen des Vorhabens nachweisen.

Dies repräsentiert einerseits eine methodische Schwäche im Projektmanagement, da es nicht gelungen ist, im Rahmen des Monitorings Beiträge zu Verbesserungen für die Zielgruppe zu beobachten und nachzuweisen. Andererseits wurden die erwarteten Wirkungen auf die eigentliche Zielgruppe der Vorhaben in den Evaluierungen sehr oft als zu ehrgeizig beurteilt, speziell unter Berücksichtigung der Kontexte, in welchen die Vorhaben agieren.

Bildungsqualität

Bezüglich der Bildungsqualität wird deutlich, dass diese in den evaluierten Projekten durchweg mit guten Schulleistungen in Verbindung gebracht wird. Als Messgrößen werden Schulnoten, Primarschul-Abschlussquoten und Übertrittsquoten in die Sekundarschule oder weiterführende Ausbildungsgänge herangezogen. Nicht in allen Fällen kann jedoch von guten Noten und erfolgreichen Abschlüssen darauf geschlossen werden, dass die betreffenden Schüler/-innen Kompetenzen aufbauen, die sie später nutzbringend verwenden können. In mehreren Evaluierungen werden Hinweise zur Bedeutung der Relevanz von Bildungsinhalten für die jeweiligen Schüler/-innen in ihrem Lebenskontext sowie in Bezug auf eine spätere Anschlussfähigkeit für die berufliche Aus- und Weiterbildung gemacht. Andererseits spiegelt sich darin die Reichweite der evaluierten Projekte wider, deren Ziele – gemäß der auftraggeberseitigen Vorgaben – auf der Ebene der Mittler (v.a. Bildungsministerien, Schulämter, Lehrpersonen) angesiedelt waren und Projekte nur indirekte Beiträge für Schüler leisteten.

Systemrelevanz

In Bezug auf die Systemrelevanz zeigt sich, dass obwohl die Relevanz der Vorhaben gemäß DAC-Kriterium sehr positiv beurteilt wurde, diese oft keine nachhaltige Wirkung auf der Systemebene entfaltet, wenngleich es hier auch positive Beispiele gab. In einigen Fällen war zwar das individuelle Capacity Building erfolgreich, der geplante Transfer in das System scheiterte jedoch. Die Vorhaben operieren alle in einem mehr oder weniger fragilen und unberechenbaren Kontext, was eine Wirkung auf Systemebene erheblich erschwert. Absichtserklärungen seitens der Regierungspartner bezüglich Reformen finden oft keinen Niederschlag in einer entsprechenden Umsetzung. Vielfach wurde der Verlauf von Reformprozessen von Vorhaben bei der Planung und auch während der Umsetzung zu optimistisch eingeschätzt. Motivation und Bedürfnisse der verschiedenen, in die Reformprozesse involvierten Akteure wurden bei der Planung nicht ausreichend analysiert. Stellvertretend scheint uns diesbezüglich das Beispiel in Ghana, wo die Regierungspartner bei der Umstellung von einem autoritätsbasierten zu einem informationsbasierten Ansatz um ihren Einfluss fürchteten. Des Weiteren wird die Laufzeit des Vorhabens einige Male kritisch thematisiert. Beispielsweise wird in Mosambik darauf hingewiesen, dass es für eine Wirkung auf systemischer Ebene eine langjährige und intensive Zusammenarbeit brauche. Es gab auch Hinweise, wo eine ungenügende Priorisierung und Politikwechsel seitens der BMZ die Systemrelevanz verhinderten oder zumindest gefährdeten. Im Vergleich zu den UEs zeichnen die PFKs – die einen Ausschnitt aus dem aktuellen Portfolio der GIZ im Bildungsbereich wiedergeben – ein etwas positiveres Bild bezüglich der Systemrelevanz.

Leistungsmix

Die Bildungsprojekte, die in der Synthese untersucht wurden, bedienen sich unterschiedlicher methodischer Ansätze, in denen Leistungen verschieden kombiniert werden. Die Erbringung der einzelnen Leistungen wird meist erfolgreich bewertet. Dabei zeigte sich, dass auch sehr vielversprechende Ansätze und Leistungen häufig erst im Zusammenspiel mit anderen Maßnahmen ihre Wirkung entfalten konnten. Besonders erfolgversprechend erwiesen sich beispielsweise die Verbindung von Lehrerfortbildungen mit der Entwicklung und Einführung neuer Lehrmittel sowie die Verwendung eines Mehrebenenansatzes. Es gilt festzuhalten, dass für die Wirkungserzielung auf Mittler- und Zielgruppenebene der kontextspezifisch adäquate Mix der einzelnen Leistungen von zentraler Bedeutung ist. Dabei erwies es sich als für das Engagement der Partner förderlich, wenn es dem Vorhaben gelang, an Bestehendes anzuknüpfen, wie beispielsweise im Tschad an die Tradition von Elternvereinen.

Geberkoordination

Eine Geberkoordination mit anderen bilateralen sowie multilateralen Geldgebern findet in der Mehrheit der Vorhaben statt. Ein zentrales Anliegen ist evaluierungstypübergreifend der Wunsch nach einer Verstärkung der Geberkoordination, um die in gewissen Fällen entstandenen Parallelstrukturen beziehungsweise Überschneidungen in Zukunft vermeiden zu können.

Hinsichtlich der Kooperation zwischen TZ und FZ lässt sich für UEs und PFKs feststellen, dass in den meisten Vorhaben eine solche Kooperation stattfand bzw. immer noch stattfindet. In den beiden dezentralen Evaluierungen spielt die Kooperation zwischen TZ und FZ eine geringere Rolle als in den Vorhaben der beiden anderen Evaluierungstypen. Die Qualität der Kooperation wird größtenteils als positiv bewertet.

5 Empfehlungen⁹

Manageriale Ebene

Für eine bessere **Nachhaltigkeit** des Vorhabens und eine realistische Einschätzung dieser empfohlen wir folgende Maßnahmen:

1. Im Dialog mit dem Hauptauftraggeber BMZ sollte die GIZ versuchen darauf hinzuwirken, dass realistische Ziele gesetzt und eine angemessene und finanziell abgesicherte Laufzeit geplant werden. In der Regel ist für die Umsetzung und die Konsolidierung von Reformprozessen eine **Langzeitorientierung und Konstanz** seitens des Vorhabens sinnvoll.
2. Für die Vorbereitung und Planung von Bildungsvorhaben sollte die GIZ prüfen, inwiefern entsprechende Instrumente im Auftragsmanagement – wie die Erarbeitung von **Akteurslandkarten** und die Durchführung von **Capacity Assessments** – noch konsequenter und gewinnbringender eingesetzt werden können. Zu klären wären dabei die Identifikation relevanter Stakeholder, eine Analyse von deren **Ownership und Motivation** sowie deren Rolle in bildungspolitischen Reformprozessen. Damit kann das Vorhaben mögliche Widerstände und Risiken antizipieren. Für den Bildungssektor sind die relevanten Akteure typischerweise, aber nicht ausschließlich Schüler/-innen, Eltern, Lehrpersonen, Schulverwaltungen und die jeweiligen für die Bildung zuständigen Regierungsstellen von der lokalen bis zur nationalen Ebene.
3. Die GIZ sollte bereits während der Planung von Projekten **Exit- und Konsolidierungsstrategien** in Zusammenarbeit mit den lokalen Partnern und mit der nötigen Flexibilität erstellen, um auf sich verändernde Rahmenbedingungen eingehen zu können.
4. In der Durchführung von Projekten sollten noch konsequenter als bisher **partizipative Ansätze** unter **Einbezug aller relevanten Akteure** (wie zum Beispiel die Einbeziehung von Eltern oder die Entwicklung von Unterrichtsmaterial unter Einbezug von Personal aus Bildungsverwaltungen, Lehrerbildungsinstitutionen und von Lehrern) verwendet werden, die zwar kosten- und zeitintensiv sind, aber qualitativ hochstehende und nachhaltige Resultate fördern. Gleichzeitig unterstützen solche partizipative Ansätze die gegenseitige Wertschätzung und das Vertrauen zwischen der GIZ und ihren Partnern.
5. In Fällen, in denen Reformprozesse mit staatlichen Partnern nicht vorangetrieben werden können, sollte die GIZ in ihren Projekten verstärkt mit **zivilgesellschaftlichen Nichtregierungspartnern** zusammenarbeiten. Im Bildungssektor können NROs durch Mitwirkung bei Konzeption und Finanzierung von Bildungsprojekten vitale Funktionen übernehmen. Zivilgesellschaftliche Organisationen besitzen oft eine größere Flexibilität und Veränderungsmotivation als Regierungsstellen, haben andererseits jedoch weniger direkten Einfluss auf die Systemebene. In den von uns ausgewerteten Evaluierungsberichten, in denen NROs thematisiert wurden, wurden ihre Dynamik und Motivation positiv beurteilt. Auch diese Empfehlung ist jedoch kontextabhängig. So gilt es in Gebieten mit regierungsfreundlichen

⁹ Eine Gegenüberstellung von Empfehlungen sowie der Basis der Empfehlungen befindet sich im Anhang auf Seite 89.

sowie regierungsfeindlichen NROs im Vorfeld mögliche Nebeneffekte einer Zusammenarbeit zu bedenken.

Hinsichtlich der **M+E Systeme, insbesondere in Bildungsprojekten**, werden folgende Schritte empfohlen:

6. M+E Systeme müssen verbessert werden, um eine fortlaufende Beobachtung des Projektverlaufs und der Steuerungsentscheidungen, sowie einen **Nachweis über erzielte Wirkungen** zu ermöglichen. Für den Bildungsbereich sollte die GIZ Energie in die Erarbeitung von Lösungen zur Messung von Beiträgen auf der Zielgruppenebene investieren, auch wenn diese in der Planung der Vorhaben außerhalb des Verantwortungsbereichs der GIZ liegt. In diesem Zusammenhang sollte die Sinnhaftigkeit von Indikatoren immer wieder überprüft werden. Im GIZ-Kontext sind Indikatoren sinnvoll, die über Veränderungen auf der Ebene von Mittlern Schlüsse auf Verbesserungen der Lernleistungen bei Schülerinnen und Schülern zulassen. Auf Zielgruppenebene bieten sich besonders regionale Lernstandsmessungen wie beispielsweise SACMEQ oder PASEC an, welche den regionalen Kontexten gerecht werden.
7. Innovationen sollten auf ihre breite Anwendbarkeit und eine weitere Verbreitung im Bildungssektor anderer Länder überprüft werden, der Transfer von guten Beispielen mit Anpassungen für den anderen Kontext kann signifikante Synergien bewirken. Die bestehenden **M+E Systeme** sind nicht ausreichend geeignet dafür, die **Kontextfaktoren**, unter denen bestimmte Ansätze ihre Wirkungen entfalten können, empirisch zu erfassen. Dies wäre aber nötig, um die Übertragbarkeit von Innovationen realistisch einschätzen zu können. Dafür wird eine wissenschaftliche Begleitforschung empfohlen. Eine Verbindung des M+E Systems mit der Akteurslandkarte wäre zu prüfen.

Sektorspezifische bzw. fachliche Ebene:

8. Wenn in einem Partnerland **Bildungsqualität mit Lernleistungen** und diese wiederum mit der Kompetenz der Schüler gleichgesetzt wird, sollte die GIZ in der Beratung ihrer Partner im Bildungswesen thematisieren, woran Lernleistungen sinnvoll gemessen werden können. Vorsicht ist geboten, wenn dazu ausschließlich Schulnoten herangezogen werden. Die Aussagekraft von Schulnoten ist in der Bildungsforschung immer wieder Gegenstand von Debatten, darüber hinaus spielt es eine wichtige Rolle, wie Schulnoten zustande kommen. Ob Schüler/-innen in der Lage sind ihr Wissen nutzbringend anzuwenden und ob ihre Kompetenzen den Anschluss an die berufliche Aus- und Weiterbildung ermöglichen, ist somit nicht automatisch anhand von Schulnoten abzulesen. Kompetenzbasierte Lehrpläne, Standards sowie Prüfungen können ein gangbarer Weg der Verbesserung sein; ein Beispiel für eine solche Vorgehensweise stellen die Bildungsstandards in Deutschland dar (vgl. z.B. Klieme, 2007; Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2004). Die Übertragbarkeit von Bildungsstandards in EZ-Kontexte stellt mit Sicherheit eine enorme Herausforderung mit vielen Unbekannten dar. Uns sind diesbezüglich keine Erfahrungen bekannt. Ein eng verwandtes Thema ist das so genannte „Policy borrowing“ bzw. der sogenannte „Policy transfer“, bei dem zwischen Nationen ein Austausch an (Bildungs-)konzepten und Herangehensweisen erfolgt (z.B. Verfahren der Leistungsmessung auf Schulsystemebene). Hierzu gibt es begrenzte Erfahrungen (z.B. Cross, Mungadi & Rouhani, 2002; Johnson, 2006, weitere Referenzen sind im Literaturverzeichnis

im Anhang zu finden), die den vorsichtigen Schluss zulassen, dass das „Kopieren“ von bildungspolitischen Konzepten zwar durchaus gelingen kann, der Vorgang in seiner Komplexität jedoch oft unterschätzt wird. Häufiger kommt der Fall vor, dass ein Konzept sich in einer Lokalität als erfolgreich erweist und daraufhin in eine andere Lokalität mit geringfügigen Anpassungen verpflanzt wird. Dabei werden Kontextfaktoren wie mögliche Kulturunterschiede sowie die Entstehungsgeschichte eines bildungspolitischen Konzepts und die dahinter stehenden motivationalen Wirkungen oft außer Acht gelassen. In der neuen Lokalität scheitert das zuvor erfolgreiche Bildungsprojekt dann in vielen Fällen.

Schließlich wurde von der OECD im Jahr 2013 „Pisa for Development“ gestartet (<http://www.oecd.org/pisa/aboutpisa/pisafordevelopment.htm>). Hier geht es vereinfacht darum, PISA so zu erweitern, dass es für Entwicklungsländer relevanter bzw. „passender“ ist, gleichzeitig dabei die Vergleichbarkeit mit den anderen PISA-Nationen zu wahren. Diese Entwicklung selbst stellt keine Bildungsstandards dar, kann die Debatte um Standards aber in vielen Ländern befeuern (wenn die Entwicklung in Folge von PISA ähnlich wie seinerzeit in einigen Industrienationen verläuft). Somit erscheinen uns Bildungsstandards als ein Weg hin zum kompetenzbasierten Lernen und als ein Thema mit Zukunftspotenzial in Entwicklungskontexten (vorbehaltlich einer Kontextualisierung in den Partnerländern der deutschen Entwicklungszusammenarbeit).

9. **Zugang zu Bildung und Bildungsqualität** stehen meist in einem engen Zusammenhang. Wir empfehlen beide Anliegen in Bildungsvorhaben weiterhin konsequent im System und nicht sequenziell im Blick zu haben.
10. Sofern der jeweils lokale Kontext einen entsprechenden Rückhalt dafür bietet, empfehlen wir weiterhin in der Beratung von Partnern im Bildungswesen auf eine Veränderung **von stark lehrerzentrierten hin zu mehr schülerzentrierten Unterrichtsformen** hinzuwirken, insbesondere in Kombination und Ergänzung mit (direktiven) Orientierungen und Ordnungsstrukturen (vgl. Hattie, 2009). Damit dieses Potenzial wirksam wird, sind insbesondere folgende Faktoren bei der Lehrerfortbildung und der Entwicklung von Lehrmaterialien zu berücksichtigen:
 - Ein Wechsel von lehrerzentrierten hin zu schülerzentrierten Unterrichtsformen bedingt tiefgreifende Veränderungen der Haltung und des Selbstverständnisses von Lehrern, was wiederum eng mit Ownership-Fragen verknüpft ist.
 - Qualitativ guter schülerzentrierter Unterricht benötigt noch besser ausgebildete Lehrkräfte als ein lehrerzentriertes Setting.
 - Erfolgreiche schülerzentrierte Unterrichtsformen benötigen entsprechend angepasste Lehrpläne, Unterrichtsmaterialien und Prüfungssysteme.
11. **Unterrichtsmaterial** sollte weiterhin nach Möglichkeit partizipativ und in hoher Qualität entwickelt werden, mit Erprobungsschlaufen in der Praxis (Mikroebene), gekoppelt mit Lehrerbildung und Organisationsentwicklung / Lehrplanentwicklung auf Makroebene.

6 Anhang

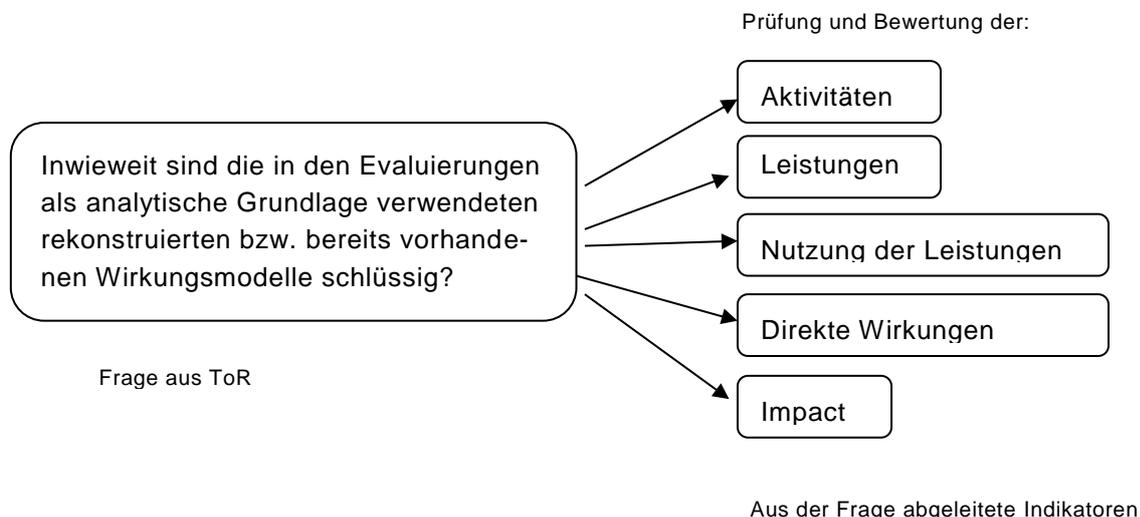
6.1 Ampelsystem

Zur Visualisierung der Bewertungen wurde ein Ampelsystem entwickelt. In diesem Ampelsystem sind einerseits jeweils alle Evaluierungen eines Typs aufgeführt (horizontale Ebene), andererseits alle Bewertungskriterien (vertikale Ebene). Dies ermöglicht es, einen schnellen Überblick über die Ergebnisse der Bewertungen zu erhalten. So werden Stärken und Schwächen der Evaluierungen wie auch Stärken und Schwächen der Erfüllung einzelner Kriterien über Evaluierungen hinweg rasch ersichtlich.

	Evaluierungen									Durchschnitt. pro Kriterium
	Eval1	Eval2	Eval3	Eval4						
Wirkungskette (max. 15 Punkte)	4	11	15	14						12
Methodische Transparenz (max. 5 Punkte)	1	3	4	5						4
Design und Datenerhebung (max. 3 Punkte)	0	2	3	2						2
Datenbasis (max. 5 Punkte)	4	4	3	4						4
DAC-Kriterien (max. 15 Punkte)	11	12	13	12						12
Empfehlungen (max. 8 Punkte)	7	8	8	8						7
Evaluierungsbewertung (max. 18 Punkte)										
	15	15	15	15	17	16	15	15	17	

Tabelle 14: Zusammenfassung der Ergebnisse der Meta-Evaluierung der UEs

Im Sinne einer größtmöglichen Transparenz wollen wir nachfolgend kurz erklären, wie die einzelnen Bewertungen zustande kommen und das Ampelsystem näher erläutern.



Basierend auf den Evaluierungsfragen aus den ToR wurden Bewertungskriterien für die einzelnen Bereiche abgeleitet. In manchen Bewertungskriterien sind mehrere Fragen aus den ToR berücksichtigt. Aus den Bewertungskriterien wurden wiederum Indikatoren abgeleitet. In manchen Fällen (z.B. beim Bewertungskriterium „Methodisches Vorgehen“) war es ausreichend, nur wenige Indikatoren abzuleiten, um das Bewertungskriterium ausreichend zu behandeln. Die folgende Tabelle stellt ein Beispiel für ein Bewertungskriterium mit wenigen Indikatoren dar:

Transparenz im methodischen Vorgehen	Bewertung
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja (was für die Evaluatoren zugänglich war)
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

In anderen Fällen (z.B. „Berücksichtigung der Wirkungsebenen“) wurde eine höhere Zahl an Indikatoren gebildet, um das Bewertungskriterium in seiner Gänze zu berücksichtigen. Die nachfolgende Tabelle stellt ein Beispiel für ein Bewertungskriterium dar, aus welchem eine größere Zahl an Indikatoren abgeleitet wurde:

Berücksichtigung der Wirkungsebenen		
Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt?¹⁰ (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Ja
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		12

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Bei der Bewertung einzelner Kriterien haben wir im Fall ungenügender Datengrundlagen oder nicht vorhandener Datengrundlagen 0 Punkte vergeben.

In einem nächsten Schritt wurden für jedes Bewertungskriterium Ampel-Beurteilungen gebildet. Diese Ampel-Beurteilungen basieren auf den erreichten Punktzahlen bei den verschiedenen Indikatoren. Beispielsweise ist aus unserer Sicht beim Kriterium „Berücksichtigung der Wirkungsebenen“ nur dann eine grüne Beurteilung (also sehr gut, keine Beanstandung) möglich, wenn von den möglichen 15 Punkten mindestens 13 erfüllt sind. Auf diese Weise wurden für alle Kriterien „Ampel-Bewertungen“ erstellt. Die individuellen Ampel-Beurteilungen wurden für die Gesamtberechnung in Punkte konvertiert: rote Ampel=1 Punkt, gelbe Ampel=2 Punkte, grüne Ampel=3 Punkte.

¹⁰ The project planning was done under the output oriented ZOPP logic (Objective Oriented Project Planning). The evaluators reconstructed a result chain where the ZOPP results and indicators are placed on the appropriate result chain levels.

Transparenz im methodischen Vorgehen	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja (was für die Evaluatoren zugänglich war)
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Die gelbe Ampel entspricht 2 Punkten bei der Berechnung der Gesamtbewertung

Diese Punkte ergaben die Rechengrundlage für die Gesamtbewertung der Einzelevaluierung. Da es 6 Bewertungskriterien gibt, sind minimal 6 Punkte (wenn alle Kriterien mit einer roten Ampel bewertet werden würden) und maximal 18 Punkte (wenn alle Kriterien mit einer grünen Ampel bewertet werden würden) in der Gesamtbewertung erreichbar.

Bewertungskriterium	Evaluierung	„Ampel“-Punkte für die Gesamtberechnung
Wirkungskette	12/15 Punkte entsprechen einer gelben Ampel.	+2 Punkte für die Gesamtberechnung
Methodische Transparenz	3/5 Punkte entsprechen einer gelben Ampel.	+2 Punkte für die Gesamtberechnung
Design und Datenerhebung	3/3 Punkte entsprechen einer grünen Ampel.	+3 Punkte für die Gesamtberechnung
Datenbasis	4/5 Punkte entsprechen einer grünen Ampel.	+3 Punkte für die Gesamtberechnung
DAC-Kriterien	11/15 Punkte entsprechen einer gelben Ampel.	+2 Punkte für die Gesamtberechnung
Empfehlungen	7/8 Punkte entsprechen einer grünen Ampel.	+3 Punkte für die Gesamtberechnung
Gesamtberechnung		2 + 2 + 3 + 3 + 2 + 3 = 15 von 18 möglichen Punkten

Die Resultate bezüglich der weiteren Fragen, welche keine Bewertungsgrundlage für die Robustheit von Evaluierungen darstellen, jedoch informativen Charakter haben, wurden nicht in das Ampelsystem einbezogen, sondern in einer eigenen Übersicht zusammengestellt (siehe Abbildung 3).

Mit Hilfe dieser Darstellung der Ergebnisse zu den einzelnen Kriterien über alle Evaluierungen hinweg werden Regelmäßigkeiten und Auffälligkeiten schnell nachvollziehbar.

6.2 Übersicht der zu untersuchenden Evaluierungen

Übersicht der Stichprobe QSA Bildung – Stand 27. Mai 2014

Land	Projekttitel	Projektdauer	Datum der Evaluierung
PFK			
Kosovo	Capacity Development in the Basic Education Sector in Kosovo	01/2010 - 12/2018, aktuelle Phase 01/2010 to 12/2012	17-25 April 2012
Honduras	Programa de Apoyo a la Calidad de la Educación Básica en el Marco del Plan EFA-FTI - PROEFA	09/2004 – 12/2014, aktuelle Phase 09/2007 – 12/2011	26.04.2011
Kongo	Réintégration des jeunes et jeunes adultes défavorisés au Maniema et Kivu (REMAKI)	02/2005 – 07/2015, aktuelle Phase 09/2008 – 12/2011	
Mauretanien	Programme Bonne Gouvernance (PBG), und Unterstützung von Bildungseinrichtungen und Fortbildungen im Bereich Dezentralisierung	01/2011 -12/2016; aktuelle Phase 01/2011 – 12/2013	30.05.2013
Mosambik	Förderung der Grund- und Berufsbildung (Pro-Educação)	01/2003 bis 12/2015 (13 Jahre) aktuelle Phase 09/2011 -12/2012	31.03.2012
Sri Lanka	Education for Social Cohesion	10/2005 – 09/2015, aktuelle Phase 10/2009 – 09/2012	9.1.2012
Zentralasien NA	Reform der Bildungssysteme in Zentralasien	1/2009 – 12/2016, aktuelle Phase 1/2011 – 12/2013	17. Mai 2013
Afghanistan	Promotion of Basic Education	2006-2016 aktuelle Phase 05/2010-11/2014	14.10.2013
Guinea	PROPEB	-	14.02.2011
Malawi	Programme to support Basic Education in Malawi	07/2010 – 06.2018 aktuelle Phase 07/2010 - 06/2014	10.2013
Sektorvorhaben	Sektorberatung Bildung	05/2011 – 04/2014 aktuelle Phase 05/2011 – 04/2014	20.09.2013
Benin	Kapazitätsentwicklung der Schulämter (Pro-Educ)	01/2011 bis 12/2013 aktuelle Phase 01/2011 – 12/2013	23.11.2013

Dezentrale Evaluierungen

Zentralamerika	Friedenserziehung in Zentralamerika und Kolumbien – Capacity Building zur Stärkung eines regionalen Friedensnetzwerks (EduPaz)	2007 - 2011	21.05.2013
Kosovo	Empowerment and Development of Youth Structures in Kosovo – EDYK	01.11.2009 - 31.12.2013	25.09.2013

Unabhängige Evaluierungen

Ghana (Ex-Post Evaluierung)	Pedagogic for Primary Schools/Assistance to Teacher Education Programme (ASTEP)	first phase 1996 – 2001; second phase 2001 – 2005	2012
Guinea (Schlussevaluierung)	Förderung der Grundbildung		2012
Honduras, Guatemala, Peru (Schlussevaluierung)	Bildungsqualität: Capacity Building für Institutionen der Lehrerbildung		2012/2013
Indonesien (Ex-Post)	Science Education Quality Improvement Project (SEQIP)		2012/2013
Namibia (Ex-Post Evaluierung)	Grundbildung in Namibia		2012
Pakistan (Ex-Post Evaluierung)	Basic Education for Afghan Refugees		2012
Peru (Ex-Post Evaluierung)	Programm Grundbildung PROEDUCA		2012
Sektorvorhaben (Schlussevaluierung)	„Innovative Ansätze in der schulischen und außerschulischen Bildung / Sektorberatung und Mädchenförderung in der Bildung“		2012/2013
Tschad (Ex-Post Evaluierung)	Förderung der Grundbildung		2012
Zentralamerika (Schlussevaluierung)	Öffentliches Budgetmanagement insbesondere im Bildungssektor		2012

6.3 Evaluierungsfragen der ToR

Meta-Evaluierung

M1	Inwieweit sind die in den Evaluierungen als analytische Grundlage verwendeten rekonstruierten bzw. bereits vorhandenen Wirkungsmodelle schlüssig? Inwieweit ist es gelungen, die unterschiedlichen Wirkungslogiken der GIZ Vorgängerorganisationen angemessen zu berücksichtigen?
M2	Inwieweit sind die Wirkungsebenen vergleichbar? (Nehmen die Vorhaben den gesamten Sektor in den Blick oder nur Teilaspekte?)
M3	Inwieweit wurden in den Evaluierungen Daten zu qualitativ hochwertigen Indikatoren (SMART-Kriterien) ausgewertet und/oder erhoben?
M4	Inwieweit basierten die Evaluierungsergebnisse auf verlässlichen Sekundärdaten (z.B. aus Monitoringsystemen von Vorhaben und/oder Partnern)?
M5	Inwieweit basierten die Evaluierungsergebnisse auf, im Rahmen der Evaluierungen durchgeführten Feldphasen erhobenen, Primärdaten?
M6	Wurde die methodische Vorgehensweise transparent dargestellt?
M7	Waren Designs und Datenerhebung adäquat auf den Untersuchungsgegenstand zugeschnitten?
M8	Wurden Vergleichsgruppen und Vorher-Nachher-Vergleiche berücksichtigt?
M9	Inwieweit sind die Bewertungen des Erfolgs der Vorhaben gemäß der OECD-DAC Evaluierungskriterien und des GIZ-Benotungssystems nachvollziehbar und intersubjektiv nachprüfbar? Erfolgte entsprechend die Notengebung bei den DAC Kriterien einheitlich, bzw. wird in einigen Evaluierungen deutlich strenger oder weniger streng benotet?
M10	Inwieweit sollten Einzelevaluierungen aus der Synthese ausgeschlossen werden, weil deren Ergebnisse als nicht ausreichend belastbar angesehen werden müssen? Wenn ja, bei welchen Evaluierungen ist dies der Fall?
M11	Welche Empfehlungen lassen sich für das Wirkungsorientierte Monitoring (WoM) von Bildungsprogrammen der deutschen TZ treffen?

Synthese

S1	Wie stellt sich der Erfolg der evaluierten Grundbildungsvorhaben gemäß der OECD-DAC Evaluierungskriterien (Relevanz, Effektivität, Impact, Effizienz, Nachhaltigkeit) insgesamt dar (Durchschnittswerte pro Kriterium)?	
S2	Welche Schlüsse lassen sich hinsichtlich der Evaluierungskriterien für den Sektor Bildung ziehen?	
S3	Lassen sich aus den Evaluierungen verallgemeinerbare manageriale Erfolgs- und Misserfolgskriterien ableiten?	
S4	Gibt es Beispiele besonders guter und/oder schlechter Managementpraxis, die für die Planung und Durchführung von Grundbildungsvorhaben relevant sind?	
S5	Lassen sich aus den Evaluierungen verallgemeinerbare sektorspezifische, bzw. fachliche Erfolgs- und Misserfolgskriterien ableiten?	
S6	Gibt es Hinweise, wie sich der Instrumentenmix auf den Erfolg insbesondere der Nachhaltigkeit eines Vorhabens auswirkt? Kristallisieren sich positive Beispiele für einen Instrumentenmix heraus?	
S7	In wie weit wurden Querschnittsthemen (z.B. Chancengerechtigkeit, Gleichberechtigung der Geschlechter, gute Regierungsführung, Armutsbekämpfung) in den Evaluierungen thematisiert? Welche Rückschlüsse lassen sich aus den Analysen zu den Querschnittsthemen im Bereich Grundbildung ziehen?	
S8	Gibt es Beispiele besonders erfolgsversprechender methodischer Ansätze, die für die Planung und Durchführung von Grundbildungsvorhaben relevant sind?	
S9	Welche Aussagen können aufgrund der Evaluierungsergebnisse zu folgenden übergreifenden Fragen getroffen werden: ¹¹	Alignment: Sind die Maßnahmen kohärent mit internationalen und nationalen Zielvorgaben und Strategien im Bildungssektor (z.B. MDG2, Education for All, nationale Bildungspolitik)?
		Zugang zu Bildung vs. Bildungsqualität: Zu welcher der Zieldimensionen leisten die Maßnahmen einen Beitrag?
		Systemrelevanz: Welchen Beitrag leisten die Maßnahmen zur Stärkung von Bildungssystemen?
		Leistungs-Mix: Wie ist der relative Beitrag von Leistungen zur Zielerreichung?
S10	Welche übergreifenden Empfehlungen lassen sich für die GIZ aus den Evaluierungsergebnissen ableiten für die Planung und Durchführung sowie das wirkungsorientierte Monitoring von Grundbildungsvorhaben (inkl. hinsichtlich des Einsatzes von TZ-Instrumenten und Ansätzen wie Fachkräfte, Beratung, HCD, Finanzierungen, Sachgüter und hinsichtlich der zukünftigen Integrierten Programmierung/Instrumenten-Mix)? Welche Empfehlungen lassen sich für die zukünftige Planung mit Wirkungsmodell/-matrix ableiten?	

¹¹ Die übergreifenden Fragen wurden durch den Auftraggeber aktualisiert. Diese Aktualisierung ist im überarbeiteten Analyseraster Synthese umgesetzt.

6.4 Analyseraster Meta-Evaluierung

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten		
Leistungen (Outputs)		
Nutzung der Leistungen (intermediate Outcome)		
direkte Wirkungen (Outcome)		
Impact		
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	
Die Wirkungshypothesen werden in der Evaluierung formuliert*	ja = 1 Punkt	
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig*	ja = 1 Punkt	
Gesamtbeurteilung des Bewertungskriteriums		<i>(Zahl und Farbe)</i>

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte*

*Die beiden Indikatoren zu Wirkungshypothesen wurden bei der Untersuchung der PFKs nicht berücksichtigt. Entsprechend wurde die Punkteskala bei der Gesamtbeurteilung von PFKs angepasst: Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11-13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	
Gesamtbeurteilung des Bewertungskriteriums	<i>(Zahl und Farbe)</i>

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Me-so-, Makroebene) des evaluierten Vorhabens.	
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	
Gesamtbeurteilung des Bewertungskriteriums	<i>(Zahl u. Farbe)</i>

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium			
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.			
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.			
Vergleichsgruppen wurden berücksichtigt.			
Vorher-Nachher-Vergleiche wurden berücksichtigt.			
Gesamtbeurteilung des Bewertungskriteriums			<i>(Zahl u. Farbe)</i>

Maximal 8 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz			
Effektivität			
Impact			
Effizienz			
Nachhaltigkeit			
Gesamtbeurteilung des Bewertungskriteriums			<i>(Zahl und Farbe)</i>

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	
Die Empfehlungen sind klar adressiert.	
Die Empfehlungen sind spezifisch.	
Gesamtbeurteilung des Bewertungskriteriums	<i>(Zahl und Farbe)</i>

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	(Zahl und Farbe)	
1. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	(Zahl und Farbe)
	2.2 Angemessenheit von Design und Datenerhebung	(Zahl und Farbe)
	2.3 Adäquate Datenbasis	(Zahl und Farbe)
3. Verwendung der OECD-DAC Kriterien	(Zahl und Farbe)	
4. Empfehlungen	(Zahl und Farbe)	
Gesamtbewertung der Evaluierung	(Zahl und Farbe)	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 – 12

Gelb: Durchschnittsbewertung 13 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit (M2)

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.						
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.						
Im evaluierten Vorhaben ist Bildung eine Subkomponente.						
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.						
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	

6.5 Analyseraster Synthese

1. Manageriale Erfolgs- und Misserfolgskfaktoren (S3, S4)

Aussagen zu guter Managementpraxis	
Aussagen zu schlechter Managementpraxis	

Capacity Works Erfolgsfaktoren	Aussagen zu Erfolgsfaktoren	Aussagen zu Misserfolgskfaktoren
Strategie (klare und plausible strategische Orientierung)		
Kooperation (Klarheit darüber, mit wem und wie kooperiert wird)		
Steuerungsstruktur (funktionsfähige Steuerungsstruktur)		
Prozesse (Klarheit über die strategisch wichtigen Prozesse)		
Lernen und Innovation (Maßnahmen, um Lernkompetenz aufzubauen und zu sichern)		

2. Sektorspezifische bzw. fachliche Erfolgs- und Misserfolgskfaktoren (S5)

Aussagen zu Erfolgsbeispielen	
Aussagen zu Misserfolgskbeispiele	

3. Instrumentenmix (S6)

Hinweise auf Wirkungsweise von Instrumentenmix (insb. auf Nachhaltigkeit)	
Positive Beispiele für Instrumentenmix (insb. auf Nachhaltigkeit)	

4. Querschnittsthemen (S7)

Querschnittsthemen	Thematisiert (ja/nein)	Falls ja, welche Aussagen
Konfliktbearbeitung / Konfliktprävention		
Geschlechtergleichberechtigung		

5. Übergreifende Fragen (S9)

Übergreifende Fragen	Spezifische Fragen	Antworten auf spezifische Fragen
Wirkungsnachweise	Welche Rückschlüsse können in Bezug auf belegbare Wirkungen auf Zielgruppenebene gemacht werden?	
Bildungsqualität	Welche Rückschlüsse können gezogen werden, wie Bildungsqualität verstanden wird?	
	Welche Rückschlüsse können gemacht werden in Bezug darauf, welche Instrumente zur Verbesserung der Bildungsqualität eingesetzt werden?	
Systemrelevanz	Welche Rückschlüsse können zur dieser Frage gezogen werden: „Gibt es in den Vorhaben auch Maßnahmen / Prozesse/Instrumente, die darauf ausgerichtet sind, nicht nur einzelne Schwachstelle im Bildungssystem beheben (z.B. defizitäre Struktur der Lehrerbildung), sondern Reformprozesse des gesamten Bildungssystems zu unterstützen (hinsichtlich Organisationsentwicklung und Verwaltungsreformen in beteiligten Ministerien und auf dezentraler Ebene)?“	
Krisenprävention und Konfliktbearbeitung	Welche Rückschlüsse lassen sich in Bezug auf diese Frage ziehen: „Welche stabilisierenden bzw. konfliktmindernden oder destabilisierenden bzw. konfliktverschärfenden Wirkungen hat(te) das Vorhaben auf gesamtgesellschaftlicher Ebene?“	
	Welche Rückschlüsse lassen sich in Bezug auf diese Frage ziehen: „Inwieweit wurden bei der Vorbereitung und Durchführung des Vorhabens krisenpräventive und friedensfördernde Maßnahmen berücksichtigt?“	
Leistungsmix	Welche Rückschlüsse können bzgl. der relativen Beiträge der Leistungen gezogen werden?	
Geberkoordination	Welche Rückschlüsse können in Bezug auf die Geberkoordination gezogen werden?	
	Welche Rückschlüsse können in Bezug auf die Kooperation zwischen deutscher TZ und FZ abgeleitet werden?	

6.6 Benotung der DAC-Kriterien und der Beurteilung der Angemessenheit der Benotung

UE	GHA		GUI		HGP		INA		NAM		PAK		PER		CHA		ZA		SV		Angemessenheit der Beurteilung
	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	
Relevanz	ZM 1	ANG 2	ANG 3	ANG 1	ZM 2	ANG 2	ANG 2	ANG 2	ZM 4	ANG 3	ANG 2	ANG 2	ZM 3	ANG 3	ANG 2	ANG 2	ANG 2	ANG 1	ANG 2	ANG 1	ANG 1.7
Effektivität	ZM 2	ZM 2	ANG 2	ZM 4	ANG 2	ANG 3	ANG 3	ANG 3	ZM 4	ANG 3	ANG 3	ZM 3	ANG 3	ANG 3	ANG 3	ANG 3	ZM 2	ZM 2	ZM 2	ZM 2	ANG 2.4
Impact	ZM 2	ZM 3	ANG 3	ZM 4	ANG 3	ANG 3	ANG 3	ANG 3	ZM 4	ANG 3	ANG 3	ZM 3	ANG 3	ANG 3	ANG 3	ZM 2	ANG 2.9				
Effizienz	ANG 2	ANG 3	ANG 2	ANG 4	ANG 2	ANG 3	ANG 3	ANG 3	ANG 4	ANG 2	ANG 3	ANG 2	ANG 2	ANG 2	ZM 2	ANG 2	ANG 2	ANG 4	ZS	ZS	ANG 2.6
Nachhaltigkeit	ANG 2	ANG 3	ANG 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ZM 3	ANG 2.8

PFK	BEN		GUI		MAW		MTN		MOZ		SV		ZAS		Angemessenheit der Beurteilung
	A	B	A	B	A	B	A	B	A	B	A	B	A	B	
Relevanz	1	ANG 1	ANG 1	ANG 1	ANG 2	ANG 2	ANG 1	ANG 1	ZM 2	ANG 1	ANG 2	ANG 1	ANG 1	ANG 1	ANG 1.4
Effektivität	3	ZM 2	ZM 2	ZM 2	ZM 3	ANG 2	ZM 3	ANG 2	ANG 3	ANG 3	ANG 2	ANG 2	ANG 2	ANG 2	ANG 2.4
Impact	2	ZM 3	ZM 3	ZM 3	ZM 3	ANG 3	ZM 4	ANG 3	ANG 3	ANG 3	ANG 2	ANG 2	ANG 2	ANG 2	ANG 2.9
Effizienz	3	ANG 2	ANG 2	ANG 4	ANG 4	ZS	ANG 2	ANG 2	ANG 3	ANG 2.9					
Nachhaltigkeit	3	ANG 3	ANG 3	ZM 4	ANG 3	ANG 3	ZM 4	ANG 3	ANG 3	ANG 3	ANG 2	ANG 2	ANG 2	ANG 2	ANG 3.1

Benotung DAC Kriterium PFK und UE-übergreifend	Ø
Relevanz	1.6
Effektivität	2.4
Impact	2.9
Effizienz	2.7
Nachhaltigkeit	2.9

UES	GHA	GUI	HGP	INA	NAM	PAK	PER	CHA	ZA	SV	Ø
Benotung Gesamtbewertung	2	3	2	4	2	3	3	3	3	2	2.6

PFKS	BEN	GUI	HON	COD	MAW	MTN	MOZ	SV	ZAS	Ø
Benotung Gesamtbewertung	2	2	3	3	3	3	4	3	2	2.7

Benotung Gesamtbewertung PFK- und UE-übergreifend	Ø
	2.7

6.7 Übersicht über Programmenden in den einzelnen Vorhaben der UEs

Ghana	Guinea	Indonesien	Namibia	Pakistan
<p>Phase wurde wie vorgesehen beendet.</p> <p>Es gab keine wirkliche Konsolidierungs- und keine explizite Exitstrategie</p>	<p>Wegen des Militärputsches und auf Weisung des BMZ wurde das Vorhaben 2012 abgebrochen bei einer ursprünglichen Laufzeit bis 2015.</p> <p>Eine Exitstrategie fehlte.</p>	<p>Phase wurde wie vorgesehen beendet.</p> <p>Keine Angaben zu einer Exitstrategie.</p>	<p>Phase wurde wie vorgesehen beendet.</p> <p>Eine Exitstrategie war nicht dokumentiert, Vorhaben kommunizierte aber klar das Ende der Förderung.</p>	<p>BMZ entschied 2002 die Finanzierung nicht fortzusetzen.</p> <p>Vorhaben lief bis 2007, jedoch mit einer Unsicherheit bezüglich der Finanzierung.</p> <p>Eine Exitstrategie wurde erst diskutiert nachdem Entscheid der BMZ seine Finanzierung nicht fortzusetzen.</p>
Honduras, Guatemala, Peru	Peru	Zentralamerika	Tschad	
<p>Wegen Putsch in Honduras wurde Vorhaben in diesem Land 2009-10 ausgesetzt und vorübergehend nach Guatemala verlagert.</p> <p>Im Zuge der Fusion zur GIZ ließ die Unterstützung seitens der deutschen TZ stark nach (Ausnahme Guatemala).</p> <p>Es war keine Exitstrategie erkennbar.</p>	<p>Wegen anderer politischer Schwerpunktsetzung des BMZ wurde Laufzeit um zwei Jahre verkürzt mit nur einem Jahr Vorankündigung.</p> <p>Eine Exitstrategie musste unter hohem Zeitdruck erarbeitet werden, was kaum Gelegenheit für eine Konsolidierungs- und Überleitungsstrategie bot.</p>	<p>Bis ein Jahr vor Phasenende wurde mit einer Verlängerung für die Konsolidierung gerechnet, welche jedoch nach der Fusion zur GIZ nicht entschieden wurde.</p> <p>Konzept wurde in Richtung einer Exitstrategie überarbeitet.</p>	<p>Einer weiteren Verlängerung wurde nicht zugestimmt.</p> <p>Ab 2008 war der Tschad aufgrund mangelnder Entwicklungsorientierung der tschadischen Regierung kein deutsches Partnerland mehr.</p> <p>Eine Exitstrategie war nicht erkennbar.</p>	

6.8 Tabellen zu den Aspekten der Vergleichbarkeit (UE, PFK, DE)

		UE (9)									Gesamt
		GHA	GUI	HGP	INA	NAM	PAK	PER	CHA	ZA	
Zuordnung Sektor Bildung	Gesamter Sektor Bildung (J/N) ¹²	N	N	N	N	N	N	N	N	N	N (9)
	Teilaspekte des Sektors (welche)	J (GB)	J (GB)	J (GB)	J (GB)	J (GB)	J (BB, GB)	J (GB)	J (GB)	N	BB (1), GB (8), N (1)
	Subkomponente (welche)	N	N	N	N	N	N	N	N	J (GB)	GB (1)
	Wirkung auf Sektor Bildung (J/N)	-	-	-	-	-	-	-	-	-	-
Region (AP/LA/NO/SA/SE/ZAs)		SA	SA	LA	AP	SA	NO	LA	SA	LA	AP (1), LA (3), NO (1), SA (4)
Projekttyp (B/R, SV)		B	B	R	B	B	B	B	B	R	B (7), R(2)
Zeitraum (vor/nach Fusion)		V	V	V/N	V	V	V	V	V	V/N	V (7), V/N (2)
Durchführungsorganisation (GTZ/InWEnt/GIZ)		GTZ	GTZ	In-WEnt /GIZ	GTZ	GTZ	GTZ	GTZ	GTZ	In-WEnt /GIZ	GTZ (5), InWEnt (1), InWEnt/GIZ (2)

¹² Abkürzungen für diese Tabelle: J/N – Ja/Nein; BB – Berufsbildung, GB – Grundbildung; FP – Friedenspädagogik; GR-Gute Regierungsführung; JF – Jugendförderung; AP – Asien/Pazifik; LA – Lateinamerika; NO – Naher Osten; SA – Sub-Sahara Afrika; SE Südosteuropa; ZAs – Zentralasien; B – bilaterales Vorhaben; R – regionales Vorhaben; SV - Sektorvorhaben.

		PFK (12)												Gesamt
		AFG	BEN	GUI	HON	COD	KOS	MAW	MTN	MOZ	SRI	SV	ZAs	
Zuordnung Sektor Bildung	Gesamter Sektor Bildung (J/N)	N	N	N	N	N	N	N	N	N	N	J	N	N (11), J (1)
	Teilaspekte des Sek- tors (wel- che)	J (GB)	N	J (GB)	J (GB)	J (GB, BB)	N	J (GB)	N	J (GB, BB)	J (GB)	-	N	BB (2), GB (7),
	Subkom- ponente (welche)	N	N	N	N	N	J (GB)	N	N	N	N	-	J (GB, BB)	BB (1), GB (2)
	Wirkung auf Sektor Bildung (J/N)	-	J (GB)	-	-	-	-	-	J (GR)	-	-	-	-	J (2)
Region (AP/LA/NO/SE/SA/Z A)		NO	SA	SA	LA	SA	SE	SA	SA	SA	AP	-	ZAs	AP (1), LA (1), NO (1), SE (1), SA (6), ZA (1)
Projekttyp (B/R, SV)		B	B	B	B	B	B	B	B	B	B	SV	R	B (10), R (1), SV (1)
Zeitraum (vor/nach Fusion)		V/N	V/N	V	V/N	V/N	V/N	V/N	V/N	V/N	V/N	N	V/N	N (2), V (1), V/N (9)
Durchführungsorga- nisation (GTZ/InWEEnt/GIZ)		GTZ/ GIZ	GTZ/ GIZ	GTZ	In- WEEnt /GIZ	GTZ/ GIZ	In- WEEnt /GIZ	In- WEEnt /GIZ	GTZ/ GIZ	GTZ/ GIZ	GTZ/ GIZ	GIZ	GTZ/ GIZ	GIZ (1), GTZ (1), GTZ/GI Z (7), In- WEEnt/GI Z (3)

		DE (2)		Gesamt
		KOS	ZAM	
Zuordnung Sektor Bildung	Gesamter Sektor Bildung (J/N)	N	N	N (2)
	Teilaspekte des Sektors (welche)	N	N	N (2)
	Subkomponente (welche)	N	J (GB)	GB (1), N (1)
	Wirkung auf Sektor Bildung (J/N)	J (JF)	-	J (1)
Region (AP/LA/NO/SA/SE/ZA)		SE	LA	LA (1), SE (1)
Projekttyp (B/R, SV)		B	R	B (1), R (1)
Zeitraum (vor/nach Fusion)		V/N	V/N	V/N (2)
Durchführungsorganisation (GTZ/InWEnt/GIZ)		GTZ/GIZ	In- WEnt/GI Z	GTZ/GIZ (1), InWEnt/GIZ (1)

6.9 Tabellen zur Wirkungsmessung auf Zielgruppen- und Mittlerebene

Hinweis: Daten aus Evaluierungen von Sektorvorhaben sind in den Tabellen aufgrund der aus unserer Sicht eingeschränkten Vergleichbarkeit nicht aufgeführt. Während sich als Zielgruppe auch bei SV in der Regel Kinder und Jugendliche finden lassen, werden als Mittler Führungskräfte des BMZ oder Organisationen der deutschen EZ genannt. Da Sektorvorhaben in den meisten Fällen fernab der eigentlichen Zielgruppe agieren, ist auch die Vergleichbarkeit der Bewertung der DAC-Kriterien, insbesondere von Impact und Effektivität, aus unserer Sicht problematisch.

UE	Mittler gemäß Projektvaluierung	Nachweisbare direkte Wirkung (Effektivität)	(Primär)-ZG gemäß Projektevaluierung	Nachweisbare indirekte Wirkung (Impact)
Ghana	Subject didactics specialists from the central teacher training division, subject lecturers of the pedagogical universities, directors and inspectors from the decentralized educational administration. Weitere: TED (Teacher Education Department as part from GES -Ghana Education Service), Ministry of Education (MoE). ¹³	Keine	Primarschüler	Keine (Impactebene in der Evaluierung nicht richtig berücksichtigt)
Guinea	MEPU, DNEE, PSE, Schulverwaltungen, Lehreraus- und Fortbildungsinstitutionen, Akteure der Zivilgesellschaft. Weitere: UNICEF, GOPA, CNLS, UNESCO, USAID. ¹⁴	Keine	Kinder ab Vorschulalter, Jugendliche, besonders Mädchen	FIERE-Ansatz : verbesserter Zugang, Verbleib, Schulleistungen, Verhaltensänderung bezüglich Verhütung
Honduras, Guatemala, Peru	3. Honduras: Universidad Pedagógica Nacional Francisco Morazán (UPNFM); Guatemala: Universidad Rafael Landívar (URL); Peru: Pontificia Universidad Católica del Perú (PUCP) Weitere: Bildungsministerien von Honduras (<i>Secretaría de Educación</i>), Guatemala (<i>Ministerio de Educación</i>)	Keine	Kinder und Jugendliche, besonders aus benachteiligten Bevölkerungsgruppen	Keine

¹³ UE_Ghana_Grundbildung_Hauptbericht, S.7.

¹⁴ UE_Guinea_Grundbildung_Hauptbericht, S.13, 23.

	und Peru (<i>Ministerio de Educación</i>). ¹⁵			
Indonesien	Education experts in the Ministry, Provinces and Districts, the 178 trained consultants, 5,525 advisory teachers and the lecturers of teacher training institutions and the science teachers of primary schools. ¹⁶	24% der Schulen in Indonesien haben die Science Kits, 11.2% bessere Prüfungsergebnisse	Schüler der 4.-6. Klasse	Keine
Namibia	Ministerium für Grundbildung, Sport und Kultur bzw. ab 2005 Bildungsministerium, Regional Education Offices National Institute for Education Development. ¹⁷	Verbesserter Zugang zum muttersprachlichen Unterricht	Primarschüler, Eltern und Dorfgemeinden	Keine
Pakistan	Master trainers, field education supervisors, headmasters, teachers in formal education, and master trainers and literacy instructors in non-formal education, and head of departments of the NGO BEFAre and the members of school management committees. Weitere: Commissionerate for Afghan Refugees, Peshawar, Pakistan, UNHCR, DFID, World Bank, CIDA. ¹⁸	Zugang zu Bildung für Mädchen hat sich gesteigert, Bildungsqualität hat sich gesteigert	Afghanische Flüchtlinge: Primarschüler, Menschen im non-formalen Bildungsbereich	Formale wie non-formale Bildung erhöhen Bildungslevel und verbessern Zukunftschancen
Peru	12.394 staatliche Lehrkräfte (Lehrerausbilder und Lehrer), Weitere: Ministerio de Educación del Perú (MINEDU), regionale Bildungsverwaltungen und Lehrerbildungseinrichtungen. ¹⁹	keine	Primarschüler (nur auf Ebene der indirekten Wirkung)	verbesserte Lehrqualität und höhere Lernleistungen an Grundschulen
Tschad	Pädagogische Fachkräfte auf lokaler, regionaler und schließlich nationaler Ebene sowie regionale und nationale Zusammenschlüsse von Elternvereinen. Weitere:	Interviewpartner bekräftigen Zusammenhang zwischen besserem Zugang und verbesserten schulischen Leistungen, höhere Primarabschlussraten, ver-	Kinder (vor allem Mädchen) im Grundschulalter sowie deren Eltern bzw. Elternvereine	Einschulungsraten als auch die Übergangsraten von der Primar zur Sekundarschule sind laut statistischen Daten des tschadischen Bildungsministeriums und der UNESCO in den letzten Jahren gestie-

¹⁵ UE_Honduras_Guatemala_Peru_Bildungsqualität_Hauptbericht, S.2.

¹⁶ UE_Indonesien_SEQIP_Hauptbericht, S.12.

¹⁷ UE_Namibia_Grundbildung_Hauptbericht, S.10.

¹⁸ UE_Pakistan_Grundbildung_Hauptbericht, S.6.

¹⁹ UE_Peru_Grundbildung_Hauptbericht, S.8-9.

	Tschadisches Erziehungsministerium, KfW, Weltbank im Rahmen des nationalen Bildungsprogramms PARSET. ²⁰	bessere schulische Infrastruktur der Projektschulen		gen, Fortschritte bezüglich Geschlechtergleichstellung
Zentralamerika	<p>Dialogmaßnahmen: gehobene Fach- und Führungskräfte in relevanten am Haushaltsprozess beteiligten Institutionen und Gremien, die den politischen Prozess steuern.</p> <p>Weiterbildungsmaßnahmen: leitende Fachkräfte aus den Finanz- und Planungsabteilungen der Institutionen.</p> <p>Nationale Trainingsprogramme: Mitarbeiter aus den Bildungs- und Finanzministerien der Kreisverwaltungen oder Fachkräfte der mittleren Managementebene auf nationaler Ebene.</p> <p>Weitere: Ministerio de Educación Honduras, Ministerio de Educación Guatemala</p> <p>Instituto Centroamericano para la Administración Pública (Zentralamerikanische Institut für Öffentliche Verwaltung in Costa Rica) Instituto Centroamericano de Administración de Empresas (Zentralamerikanische Institut für Betriebswirtschaft in Costa Rica), Universidad Pedagógica Nacional Francisco Morazan (Universität in Honduras), Universidad Rafael Landívar (Universität in Guatemala).²¹</p>	keine	Schüler/innen, Eltern, Lehrer/innen, alle Bürger	Keine (Impactebene in der Evaluierung nicht richtig berücksichtigt)

PFK	Mittler gemäß Projektevaluierung	Nachweisbare direkte Wirkung (Effektivität)	(Primär)-ZG gemäß Projektevaluierung	Nachweisbare indirekte Wirkung (Impact)
Benin	Vor- und Grundbildungsministerium (MEMP), COS (Comite d'orientation et de suivi), Centre International d'Etudes Pédagogiques (CIEP), DDEMP, Staatliche	Bessere Leistungsfähigkeit und Qualität ihrer Dienstleistungen	Personal unterschiedlicher Kategorien in den Schulämtern	keine

²⁰ UE_Tschad_Grundbildung_Hauptbericht, S.10 ,19.

²¹ UE_Zentralamerika_Öffentliches Budgetmanagement insbesondere im Bildungssektor_Hauptbericht, S.6, 7, 17.

	Schulämter, CEFAL, CFPEEN, ENI. ²²			
Guinea	MEPU-EC et coordination nationale du PSE. ²³	verbesserter Zugang, Verbleib und Schulleistungen, besonders im Vergleich zu nicht Projektregionen	Primarschüler	Schulangebot für Mädchen in schwierigen Lebenssituationen wurde verbessert
Honduras	4. Secretaría de Educación en todos sus niveles, Redes Educativas, Sistema Nacional de Formación de Docente, Universidad Nacional Autónoma de Honduras, Universidad Pedagógica Nacional Francisco Morazán, Instituto Nacional de Investigación y Capacitación Educativa, Unidad de Transparencia, Consejo Nacional Anticorrupción, Instituto para el Acceso a la Información Pública, algunas ONGs. ²⁴	Keine	Schüler der 1.-9. Klasse besonders aus armen und ländlichen Familien	Keine
Kongo	ONG, prestataires de services privés, services déconcentrés et décentralisés publics. (Divisions et services d'EPSP, Affaires Sociales, Jeunesse, Petites et Moyennes Entreprises; prestataires de services) ²⁵	A Maniema, 2.152 de la totalité des 2.521 jeunes, dont 65% des filles, ont obtenu l'examen TENAFEP Jusqu'au mois de mars 2011, 3.622 sur 3.707 participants à l'apprentissage, dont 37% de filles (effectifs cumulés), ont réussi les examens finaux. Jusqu'en fin février 2011, 1 877 se sont installés en groupe comme 'micro entrepreneurs'. Ces 1 877 'jeunes entrepreneurs' sont tous issus des 3 622 jeunes certifiés professionnellement, soit un taux de 51,82%. Le pourcentage pour les filles et jeunes femmes correspond à 41,96%,	Junge Erwachsene (Definition der ZG wurde kritisiert und im Laufe des Projekts erweitert)	Konfliktprävention: La mission est assez convaincue que les mesures du projet ont eu surtout cet impact.

²² PFK_Benin_2013_final, S.6, 7, 19, 55.

²³ PFK_Guinée_PROPEB_final, S.41.

²⁴ PFK_Honduras_2011_Informefinal_ohneAnhaenge_SPANISCH, S.46.

²⁵ PFK_Kongo(RDC)_xxxx_versionfinal, S.54.

		celui des jeunes hommes à 52,75%.		
Malawi	MoEST/DBE, NGOs, Malawi Institute of Education (MIE), KfW, Gopa, CIDA. ²⁶	Wiederholungsraten gingen leicht zurück, Lernerfolge sind leicht gestiegen	Schüler im Alter von sechs bis siebzehn Jahren	keine
Mauretani- en	Ministerium für wirtschaftliche Angelegenheiten und Entwicklung (MAED), Generaldirektion für Territoriale Gebietskörperschaften (DGCT), Finanzministerium, Rechnungshof (CdC), NV, Weitere: MIDEK, ENAMJ, CNITIE, ARMP. ²⁷	Verhaltensänderungen bez. Verständnis zu lokaler Interessensvertretung ist feststellbar	Fachkräfte der staatlichen Verwaltung auf nationaler und subnationaler Ebene, Fachkräfte von Bildungseinrichtungen, Abgeordnete der Nationalversammlung	keine
Mosambik	Lehrkräfte, Schuldirektor/innen, Führungskräfte der beteiligten Ministerien, Bildungsplaner und -administratoren auf Provinz-, Distrikt- und lokaler Ebene zusätzlich für die Berufsbildung: Ausbilder, Unternehmen und Verbände. Weitere: Bildungsministerium (DPEC), (SDEJT), Nationale Direktion für Berufsbildung DINET im MinED und INEFP im Arbeitsministerium, PIREP. ²⁸	keine	Kinder im grundschulpflichtigen Alter, Eltern und Gemeinden Berufsbildung: Jugendliche	keine
Zentralasien	Tutoren und Dozenten an päd. Fachschulen und Fakultäten; Lehrkräfte an allgemeinbildenden Schulen; Schulleiter an Pilot- und Ausbildungsschulen. Weitere: Abteilungen für Internationale Zusammenarbeit, für Strategische	Zunahme der Lernleistungen der Schüler/innen ist die Zielerreichung zum Ende der Phase realistisch	Kinder und Jugendliche im Schulalter (6-18 Jahre); Lehrerstudenten an den pädagogischen Fachschulen und Fakultäten.	keine

²⁶ PFK_Malawi_2013_final, S.11, 16, 56.

²⁷ PFK_Mauretanien_2013_final, S.4 ,35.

²⁸ PFK_Mosambik_20042012-final, S.6, 9, 15 ,60.

	Entwicklung und Bildungsplanung sowie für Lehrerbildung; nationale Bildungsinstitute; Fachschulen der Lehrerbildung; Provinzbehörden; Kreis-schulämter. ²⁹			
--	---	--	--	--

Hinweis: Die PFKs Afghanistan, Kosovo und Sri Lanka wurden in der Meta-Evaluierung als nicht ausreichend robust eingeschätzt und nicht im Rahmen der Synthese ausgewertet. Aus diesem Grund wurden sie auch nicht in die obige Tabelle zur Wirkungsmessung aufgenommen.

²⁹ PFK_Zentralasien_2013_Programm_Bildung_Endfassung, S.31.

DE	Mittler gemäß Projektevaluierung	Nachweisbare direkte Wirkung (Effektivität)	(Primär)-ZG gemäß Projektevaluierung	Nachweisbare indirekte Wirkung (Impact)
Kosovo	Ministry of Culture, Youth and Sports (MCYS) and its youth department (DOY), municipal youth offices and youth councils, UNICEF, Association of Kosovo Municipalities (AKM), Ministry of Local Government Administration (MLGA), Parliamentary Group of Young Deputies, KYAC, LYAC. ³⁰	keine	Jugendliche und junge Erwachsene im Alter zwischen 15 und 24 Jahren	<p>“There is an increasing involvement of young people due to the LYAC’s “political” voice.”</p> <p>“Young people participate in the local MYAP elaboration process and influence local policies and decisions. Inclusive attitudes among young people regarding gender issues and ethnic and minority groups are increasing.” “52% of young people from the EDYK-supported municipalities suggested that there were more youth participation possibilities offered, compared to only 20% from municipalities that were not supported by EDYK.”</p>
Zentralamerika	<p>Vertreter des staatlichen Bildungssektors, Mitarbeiter der Bildungsministerien, Nichtregierungsorganisationen, Basisorganisationen, organisierte Mitglieder von Bauern- oder Frauenorganisationen, nichtorganisierte Einwohner/innen von ländlichen oder städtischen Gemeinden.</p> <p>Guatemala: Proyecto de Desarrollo Santiago (PRODESSA)</p> <p>Kolumbien: Pontificia Universidad Javeriana,</p>	„Zahlreiche Beispiele sowohl auf lokaler als auch regionaler Ebene konnten identifiziert werden, die belegen, wie die Philosophie von Educación para la Paz durch die Alumni bis zum heutigen Tage in die Fläche getra-	Bevölkerungsgruppen in den von Kriegs- und Nachkriegsphänomenen gezeichneten fünf Zielländern	Berufliche Weiterentwicklung der EduPaz Alumni ist sehr positiv.

³⁰ DE_Kosovo_EDYK_SUM-UBO_final, S.7, 11, 36, 37, 43, 44.

	<p>Instituto de Estudios Sociales y Culturales (PENSAR)</p> <p>Centro de Investigaciones y Educación Popular (CINEP)</p> <p>Weitere:</p> <p>GTZ Guatemala, Kolumbien, El Salvador</p> <p>DED Nicaragua, Guatemala</p> <p>KFW Honduras</p> <p>Paulo Freire Institut (PFI) (FU Berlin)</p> <p>IDHUCA, Escuela Normal Montes de María, CBS, Arzobispado³¹</p>	gen wird.“		
--	---	------------	--	--

³¹ DE_Kolumbien_2013_EduPaz_final, S.13-15, 53.

6.10 Übersicht der Empfehlungen und der Basis für Empfehlungen

EMPFEHLUNGEN	Basis für Empfehlung
Stärkung des Erfolgsfaktors Kooperation und Stärkung der Systemrelevanz	
<p>1. Im Dialog mit dem Hauptauftraggeber BMZ sollte die GIZ versuchen darauf hinzuwirken, dass realistische Ziele gesetzt und eine angemessenen langfristig und finanziell abgesicherte Laufzeit geplant werden.</p>	<p><i>Erfolgsfaktor Kooperation:</i> Vielfach wurde der Verlauf von Reformprozessen von Vorhaben bei der Planung und auch während der Umsetzung zu optimistisch eingeschätzt.</p>
<p>2. Für die Vorbereitung und Planung von Bildungsvorhaben sollte die GIZ prüfen, inwiefern entsprechende Instrumente im Auftragsmanagement – wie die Erarbeitung von Akteurslandkarten und die Durchführung von Capacity Assessments – noch konsequenter und gewinnbringender eingesetzt werden können. Fragen, die dabei geklärt werden sollten, sind die Identifikation relevanter Stakeholder, eine Analyse von deren Ownership und Motivation, sowie ihrer Rolle in bildungspolitischen Reformprozessen.</p>	<p><i>Erfolgsfaktor Kooperation:</i> Bemängelt wurde die Kooperation hinsichtlich Motivation und Kapazitäten der nationalen Regierungspartner. Motivation und Bedürfnisse der verschiedenen, in die Reformprozesse involvierten Akteure, wurden bei der Planung nicht ausreichend analysiert.</p> <p><i>Methodische Ansätze:</i> Vierversprechend erweist sich die gezielte Verwendung eines Mehrebenenansatzes. Dieser steht dabei in engem Zusammenhang mit der oben genannten Systemrelevanz. Bei einigen UEs wurde deutlich, dass durchaus erfolgreiche Maßnahmen auf einzelnen Ebenen nur geringe oder keine breitenwirksame Wirkung entfalten konnten, weil Faktoren auf anderen Ebenen dies nicht unterstützten.</p>
<p>3. Die GIZ sollte bereits während der Planung von Projekten Exit- und Konsolidierungsstrategien in Zusammenarbeit mit den lokalen Partnern erstellen.</p>	<p><i>Erfolgsfaktor Kooperation:</i> 5. Mit Blick auf die Programmenden wurden in den UEs aber oft eine fehlende Exit-Strategie und eine ungenügende Konsolidierung bemängelt.</p>
<p>4. In der Durchführung von Projekten sollten noch konsequenter als bisher partizipative Ansätze unter Einbezug aller relevanten Akteure verwendet werden, die zwar kosten- und zeitintensiv sind, aber qualitativ hochstehende und nachhaltige Resultate fördern.</p>	<p><i>Erfolgsfaktor Kooperation:</i> In einigen Fällen wurden nicht alle relevanten Akteure ausreichend und genügend partizipativ in die Planung einbezogen.</p> <p><i>Methodische Ansätze:</i> Dabei erwies es sich als für das Engagement der Partner förderlich, wenn es dem Vorhaben gelang, an Bestehendes anzuknüpfen.</p>

<p>5. In Fällen, in denen Reformprozesse mit staatlichen Partnern nicht vorangetrieben werden können, sollte die GIZ in ihren Vorhaben verstärkt mit zivilgesellschaftlichen Nichtregierungspartnern zusammenarbeiten.</p>	<p><i>Erfolgsfaktor Kooperation:</i> 6. Positiv beurteilt wurden die Zusammenarbeit mit Nichtregierungspartnern, ihre Motivation und ihre Dynamik.</p>
<p>7. Schritte hinsichtlich der M+E System, insbesondere in Bildungsprojekten</p>	
<p>6. M+E Systeme müssen verbessert werden, um eine fortlaufende Beobachtung des Projektverlaufs und der Steuerungsentscheidungen, sowie einen Nachweis über erzielte Wirkungen auf Zielgruppenebene zu ermöglichen.</p>	<p>8. <i>Monitoring und Evaluation:</i> Im Bereich der übergreifenden, strategisch besonders relevanten Fragen, die seitens der GIZ in den ToR formuliert werden, lässt sich in Bezug auf Wirkungsnachweise zunächst das oftmalige Fehlen der Wirkungsnachweise auf Zielgruppenebene festhalten.</p>
<p>7. Innovationen sollen auf ihre breite Anwendbarkeit und eine weitere Verbreitung im Bildungssektor anderer Länder überprüft werden, der Transfer von guten Beispielen mit Anpassungen für den anderen Kontext kann erstaunliche Synergien bewirken. Die bestehenden M+E Systeme sind nicht ausreichend geeignet dafür, die Kontextfaktoren, unter denen bestimmte methodische Ansätze ihre Wirkungen entfalten können, empirisch zu erfassen.</p>	<p><i>Erfolgsfaktor Kooperation:</i> Bemängelt wurde die Kooperation hinsichtlich Motivation und Kapazitäten der nationalen Regierungspartner. Motivation und Bedürfnisse der verschiedenen, in die Reformprozesse involvierten Akteure, wurden bei der Planung nicht ausreichend analysiert.</p>
<p>9. Empfehlungen bezüglich der fachlichen Aspekte</p>	
<p>8. Wenn in einem Partnerland Bildungsqualität mit Lernleistungen und diese wiederum mit der Kompetenz der Schüler gleichgesetzt wird, sollte die GIZ in der Beratung ihrer Partner im Bildungswesen thematisieren, woran Lernleistungen sinnvoll gemessen werden. Vorsicht ist geboten, wenn dazu Schulnoten herangezogen werden.</p>	<p><i>Monitoring und Evaluation:</i> 10. Bezüglich der Bildungsqualität wird deutlich, dass diese in den evaluierten Projekten durchweg mit guten Schulleistungen in Verbindung gebracht wird. Als Messgrößen werden Schulnoten, Primarschulabschlussquoten und Übertrittsquoten in die Sekundarschule oder weiterführende Ausbildungsgänge herangezogen. Nicht in allen Fällen kann jedoch von guten Noten und erfolgreichen Abschlüssen darauf geschlossen werden, dass die betreffenden Schüler/-innen Kompetenzen aufbauen, die sie später nutzbringend verwenden können.</p>
<p>9. Zugang zu Bildung und Bildungsquali-</p>	<p><i>Methodische Ansätze:</i></p>

<p>tät stehen meist in einem engen Zusammenhang. Wir empfehlen beide Anliegen in Bildungsvorhaben weiterhin konsequent im System und nicht sequenziell im Blick zu haben.</p>	<p>Bildungsqualität und Zugang zu Bildung sind entsprechend Anliegen, die eng miteinander verknüpft sind und weder parallel noch sequenziell, sondern nur systemisch bearbeitbar sind. Dies wurde in allen Bildungsvorhaben unter Berücksichtigung des Kontextes gut berücksichtigt.</p>
<p>10. Sofern der jeweils lokale Kontext einen entsprechenden Rückhalt dafür bietet, empfehlen wir weiterhin in der Beratung von Partnern im Bildungswesen auf eine Veränderung von stark lehrerzentrierten hin zu mehr schülerzentrierten Unterrichtsformen hinzuwirken. Damit dieses Potenzial wirksam wird, ist zu berücksichtigen, dass dieser Wechsel mit tiefgreifenden Veränderungen der Wertehaltungen von Lehrpersonen verknüpft ist.</p>	<p><i>Methodische Ansätze:</i> Im Rahmen der Aus- und Fortbildung wird auch häufig eine Verschiebung von lehrerzentriertem zu schülerzentriertem Unterricht thematisiert.</p>
<p>11. Unterrichtsmaterial sollte weiterhin partizipativ und in hoher Qualität entwickelt werden, mit Erprobungsschlaufen in der Praxis, gekoppelt mit Lehrerbildung und Organisationsentwicklung / Lehrplanelwicklung auf Makroebene.</p>	<p><i>Methodische Ansätze:</i> Die Entwicklung und Verteilung von neuen Lehr- und Lernmaterialien sowie von Unterrichtsmaterialien sind weitere wichtige Maßnahmen – im Idealfall wird beides kombiniert. So ist eher gewährleistet, dass Lehrkräfte diese Materialien anwenden können und dies im Unterricht auch tun.</p>

6.11 Datenblätter (Meta-Evaluierung)

6.11.1 Analyseraster Meta-Evaluierung UE Ghana

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Ja
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		12

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja (was für die Evaluatoren zugänglich war)
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	
3	

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja (wird erwähnt)
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja (wird erwähnt)
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte

gelb: 1 – 2 Punkte

grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Literatur, Statistiken	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Semi-strukturierte und Telefoninterviews, Fragebögen, Schülerbeurteilungen, Workshops	
Vergleichsgruppen wurden berücksichtigt.	Nein		Impossible, as NALAP influence overlaid all ASTEP results; keine VG vorhanden
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Zu Projektbeginn, Projektende, fünf Jahre danach	Nur auf Zielgruppenebene
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte

gelb: 2 – 3 Punkte

grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja (eher teilweise, viel zu BMZ relevanten Themen, wurde viel über effektivitätsrelevante Faktoren geschrieben)	Teilweise	Zu mild
Effektivität	Ja	Nein	Zu mild (Hauptziel (overall objective) konnte nicht gemessen werden => keine nachvollziehbare Bewertung; wichtige Prüffragen wurden zudem außer acht gelassen)
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			11

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		15

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
					X	
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X			
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X			

6.11.2 Analyseraster Meta-Evaluierung UE Guinea

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Nein
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein (nur Nennung, keine Analyse)
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	ja
Gesamtbeurteilung des Bewertungskriteriums		11

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Sekundärdaten aus anderen Evaluierungen	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Statistiken, leitfadengestützte Intensivinterviews, Fokusgruppeninterviews, standardisierte Befragung	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Trichteransatz	Keine brauchbaren Baseline-daten
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Nein	Zu mild
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch nur am Rande pädagogische Empfehlungen ausgesprochen)
Die Empfehlungen sind spezifisch	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		15

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
				X		
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X			
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X			

6.11.3 Analyseraster Meta-Evaluierung UE Honduras, Guatemala, Peru

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Nein	Nein
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Teilweise
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		9

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	4

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	1

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	M&E Daten, Statistiken	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Fokusgruppen, semi-strukturierte Interviews, Fernbefragungen	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		„Es würde insgesamt wenig Erkenntnisgewinn erbringen, die Ausgangssituation an dieser Stelle mit Baselinedaten zu 26 Städten (inkl. dreier Hauptstädte) oder ca. ebenso vielen Provinzen in den 3 Partnerländern zu untermauern.“ (S.6)
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Teilweise	Angemessen
Impact	Ja	Ja	Angemessen
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			13

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		3
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		15

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
			X			
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	R		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
				X		
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X	X	X	

6.11.4 Analyseraster Meta-Evaluierung UE Indonesien

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Ja
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		11

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Nein
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Statistiken, M&E Daten	Teilweise schwer zugänglich
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Fragebogen, Interview	
Vergleichsgruppen wurden berücksichtigt.	Ja	ZG/KG Vergleich	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Teilweise	Zu mild
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	2	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien	2	
4. Empfehlungen	3	
Gesamtbewertung der Evaluierung	15	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Ja</i>	<i>Grundbildung</i>
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
		X				
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
					X	
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X			
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X			

6.11.5 Analyseraster Meta-Evaluierung UE Namibia

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Ja
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		15

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja		
Vergleichsgruppen wurden berücksichtigt.	Nein		Projekt betrifft das gesamte Land
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Trichteransatz	Keine brauchbaren Baselinedaten
Gesamtbeurteilung des Bewertungskriteriums			4
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Teilweise	Zu mild
Effektivität	Ja	Ja	angemessen
Impact	Ja	Teilweise	angemessen
Effizienz	Ja	Teilweise	angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			10

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	3	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien	2	
4. Empfehlungen	3	
Gesamtbewertung der Evaluierung	17	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung	ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.	Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.	Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.	Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.	-	

Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
	X					
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
					X	
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X			
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X			

6.11.6 Analyseraster Meta-Evaluierung UE Pakistan

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Teilweise
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		12

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	4

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja

Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Nein
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Literaturstudium	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interview, Fragebogen	
Vergleichsgruppen wurden berücksichtigt.	Ja	ZG/KG Interviews und Vergleiche	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		Keine brauchbaren Baseline Daten
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Ja	Angemessen
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			14

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch kaum thematisiert)
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		3
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		16

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung	<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.	<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.	<i>Ja</i>	<i>Grundbildung Berufsbildung</i>
Im evaluierten Vorhaben ist Bildung eine Subkomponente.	<i>Nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.	-	

Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Lateinamerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südosteuropa</i>	<i>Zentralasien</i>
				X		
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
					X	
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X			
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X			

6.11.7 Analyseraster Meta-Evaluierung UE Peru

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		13

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja		
Vergleichsgruppen wurden berücksichtigt.	Ja	Durchführung qualitativer Fallstudien	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Trichteransatz	Keine brauchbaren Baselinedaten
Gesamtbeurteilung des Bewertungskriteriums			5

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Teilweise	Zu mild
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (kaum thematisiert)
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	6

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	3	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien	2	
4. Empfehlungen	2	
Gesamtbewertung der Evaluierung	15	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung	ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.	Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in	Ja	Grundbildung

den Blick.								
Im evaluierten Vorhaben ist Bildung eine Subkomponente.							<i>Nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.							-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>		
			X					
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>			
			<i>B</i>	<i>B</i>				
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>			
					X			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>			
			X					
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>			
				X				

6.11.8 Analyseraster Meta-Evaluierung UE Tschad

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Teilweise
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		12

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Politik- und Strategiepapiere, Studien, GIZ Dokumente	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interview, Fragebogen	
Vergleichsgruppen wurden berücksichtigt.	Ja	Von den Evaluatoren erstellte KG	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Trichteransatz	
Gesamtbeurteilung des Bewertungskriteriums			5
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Teilweise	Angemessen
Effizienz	Ja	Teilweise	Zu mild
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Teilweise
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		15

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
					X	
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X			
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X			

6.11.9 Analyseraster Meta-Evaluierung UE Zentralamerika

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Teilweise
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		14

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Dokumentenana-lyse	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Leitfadeninter-view, standardi-sierte Interviews Fokusgruppen Diskussion	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden be-rücksichtigt.	Ja	Trichteransatz	
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			13

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (nur indirekt, explizit nicht thematisiert)
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		3
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		3
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		17

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Nein	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Ja	Grundbildung
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
			X			
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	R		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
					X	
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
				X		

6.11.10 Analyseraster Meta-Evaluierung UE SV Innovation

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Ja
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Ja
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Ja
Gesamtbeurteilung des Bewertungskriteriums		15

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

(M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Daten- und Dokumentenanalyse	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	u.a. Interviews	
Vergleichsgruppen wurden berücksichtigt.	Ja	Kontributionsanalyse	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Trichteransatz	
Gesamtbeurteilung des Bewertungskriteriums			5
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zur Benotung führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Teilweise	Zu mild
Impact	Ja	Ja	Angemessen
Effizienz	Ja	Teilweise	Zu streng
Nachhaltigkeit	ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			13

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		3
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		3
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		17

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Ja	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Nein	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					Nein	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
					X	
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
				X		
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X			
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X			

6.11.11 Analyseraster Meta-Evaluierung PFK Afghanistan

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Nein	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		6

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Nein
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	0

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja

Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Nein		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Analyseberichte, offizielle Dokumente, Monitoringberichte	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			2

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz			
Effektivität			
Impact			
Effizienz			
Nachhaltigkeit			
Gesamtbeurteilung des Bewertungskriteriums			Nicht bewertbar

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Nein
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Nein (nicht vorhanden)
Die Empfehlungen sind spezifisch.	Nein
Die Empfehlungen sind klar adressiert.	Nein
Gesamtbeurteilung des Bewertungskriteriums	0

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		1
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		1
4. Empfehlungen		1
Gesamtbewertung der Evaluierung		9

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung				ja / nein	Welche?	
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.				Nein		
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.				Ja	Grundbildung	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.				Nein		
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.				-		
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
				X		
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
			X			
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X		X	

6.11.12 Analyseraster Meta-Evaluierung PFK Benin

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Nein	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		6

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte

gelb: 1 – 2 Punkte

grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Wird nicht näher ausgeführt	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Einzel-, Gruppeninterviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja		Baslinedaten sollen noch erhoben werden
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte

gelb: 2 – 3 Punkte

grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität**	Ja	Teilweise	Zu mild
Impact**	Ja	Teilweise	Zu mild
Effizienz	Ja	Teilweise	Angemessen
Nachhaltigkeit**	Ja	Ja	angemessen
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

** Die Kriterien Effektivität, Impact und Nachhaltigkeit werden in der Evaluation nicht separiert bewertet.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Teilweise
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Nein (nicht vorhanden)
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		1
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		2
Gesamtbewertung der Evaluierung		13

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung				ja / nein	Welche?	
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.				Nein		
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.				Nein		
Im evaluierten Vorhaben ist Bildung eine Subkomponente.				Ja	Grundbildung	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.				-		
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
			X			
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
					X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
					X	

6.11.13 Analyseraster Meta-Evaluierung PFK Guinea

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Nein
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?		
	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		7

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	1

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja

Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Zur Verfügung gestellte Dokumente	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Einzel-, Gruppeninterviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	angemessen
Effektivität	Ja	Teilweise	Zu mild
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			12

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Nein
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		2
Gesamtbewertung der Evaluierung		12

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Ja	Grundbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
			X			
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X			
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X			

6.11.14 Analyseraster Meta-Evaluierung PFK Honduras

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Nein
Nutzung der Leistungen (intermediate Outcome)	Ja	Nein
direkte Wirkungen (Outcome)	Ja	Nein
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	teilweise
Gesamtbeurteilung des Bewertungskriteriums		4

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	ja

Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	ja	The World Bank, Public Expenditure Tracking and Service Delivery, 2010) Informe de Progreso Educativo, Ferrema –PREAL, 2010) Datos de la UPEG Instituto Nacional de Estadísticas Instituto Nacional de la Mujer de Honduras Censo de población 2001	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	ja	Debate de expertos	
Vergleichsgruppen wurden berücksichtigt.	nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	nein		
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	ja	teilweise	zu mild
Effektivität	ja	teilweise	zu mild
Impact	nein	teilweise	zu mild
Effizienz	ja	ja	angemessen
Nachhaltigkeit	ja	teilweise	zu mild
Gesamtbeurteilung des Bewertungskriteriums			10

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	ja
Die Empfehlungen sind spezifisch.	teilweise
Die Empfehlungen sind klar adressiert.	teilweise
Gesamtbeurteilung des Bewertungskriteriums	6

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		1
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		2
Gesamtbewertung der Evaluierung		12

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 – 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>ja</i>	<i>Grundbildung</i>
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentral-asien</i>
			X			
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
			X			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
				X	X	

6.11.15 Analyseraster Meta-Evaluierung PFK Kongo

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Nein
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		7

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Diverse Berichte, Befragungen	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Zu mild
Impact	Nein	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			13

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (nur am Rande thematisiert)
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	teilweise
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt

gelb: 3 – 6 Punkte

grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	2	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien	3	
4. Empfehlungen	3	
Gesamtbewertung der Evaluierung	15	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung	ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.	Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.	Ja	Grundbildung Berufsbildung
Im evaluierten Vorhaben ist Bildung eine Subkomponente.	Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.	-	

Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentral-asien</i>
	X					
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
			X			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X		X	

6.11.16 Analyseraster Meta-Evaluierung PFK Kosovo

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten		
Leistungen (Outputs)		
Nutzung der Leistungen (intermediate Outcome)		
direkte Wirkungen (Outcome)		
Impact		
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	
Gesamtbeurteilung des Bewertungskriteriums		<i>Nicht bewertbar</i>

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	
Gesamtbeurteilung des Bewertungskriteriums	<i>Nicht bewertbar</i>

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	
Gesamtbeurteilung des Bewertungskriteriums	Nicht bewertbar
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium			
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.			
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.			
Vergleichsgruppen wurden berücksichtigt.			
Vorher-Nachher-Vergleiche wurden berücksichtigt.			
Gesamtbeurteilung des Bewertungskriteriums	Nicht bewertbar		
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz			
Effektivität			
Impact			
Effizienz			
Nachhaltigkeit			
Gesamtbeurteilung des Bewertungskriteriums			Nicht bewertbar

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Teilweise
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Nein
Gesamtbeurteilung des Bewertungskriteriums	4

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	1	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	1
	2.3 Adäquate Datenbasis	1
3. Verwendung der OECD-DAC Kriterien	1	
4. Empfehlungen	2	
Gesamtbewertung der Evaluierung	7	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Nein</i>	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Ja</i>	<i>Grundbildung</i>
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
					X	
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
			<i>X</i>			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
					X	

6.11.17 Analyseraster Meta-Evaluierung PFK Malawi

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten		
Leistungen (Outputs)		
Nutzung der Leistungen (intermediate Outcome)		
direkte Wirkungen (Outcome)		
Impact		
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	
Gesamtbeurteilung des Bewertungskriteriums		Nicht bewertbar

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	1

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Statistiken, Bildungsmanagementdaten	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interviews mit Stakeholdern	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		Keine Baselinedaten verfügbar
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte

gelb: 2 – 3 Punkte

grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Ja	Angemessen
Effizienz	Ja	Teilweise	Zu streng
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			14

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte

gelb: 7 – 12 Punkte

grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	6

Maximal 4 Punkte: rot: 0 – 2 Punkt

gelb: 3 – 6 Punkte

grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	1	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien	3	
4. Empfehlungen	2	
Gesamtbewertung der Evaluierung	12	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Ja</i>	<i>Grundbildung</i>
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
	X					
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
			<i>X</i>			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			<i>X</i>		<i>X</i>	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
				<i>X</i>	<i>X</i>	

6.11.18 Analyseraster Meta-Evaluierung PFK Mauretaniens

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Nein	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		4

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Diverse Referenzdokumente	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Befragungen von Akteuren im Umsetzungsprozess	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		Baselinedaten stehen jedoch zur Verfügung
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Teilweise	Angemessen
Effektivität	Ja	Teilweise	Zu mild
Impact	Ja	Teilweise	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Nein	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			9

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Nicht relevant
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Nein
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		1
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		2
4. Empfehlungen		2
Gesamtbewertung der Evaluierung		12

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Nein	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					Ja	Grundbildung
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
	X					
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischenevaluierung	Schluss-evaluierung	Ex-post Evaluierung	
			X			
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X		X	

**6.11.19 Analyseraster Meta-Evaluierung PFK Mosambik
Bewertungskriterien**

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		10

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Schlussberichte von Projektphasen, Fachberichte aus Kurzeiteinsätzen, Mid Term Report der Weltbank für den Teil Berufsbildung, Evaluierungs- und Strategiedokumente des PBA, Produkte des Vorhabens	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja		
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Teilweise	Zu mild
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Ja	Angemessen
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			14

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Nein
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	2	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	2
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien	3	
4. Empfehlungen	2	
Gesamtbewertung der Evaluierung	15	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Ja</i>	<i>Grundbildung Berufsbildung</i>
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Nein</i>	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein- amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost- europa</i>	<i>Zentralasien</i>
	X					
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen- evaluierung</i>	<i>Schluss- evaluierung</i>	<i>Ex-post Evaluierung</i>	
			<i>X</i>			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X	X	X	

6.11.20 Analyseraster Meta-Evaluierung PFK Sri Lanka

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Gesamtbeurteilung des Bewertungskriteriums	<i>Nicht vorhanden</i>
Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte	

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	1
Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte	

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Nein
Gesamtbeurteilung des Bewertungskriteriums	2
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Impact Reports, Impact Monitoring Plans	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Experteninterviews, Fokusgruppeninterviews	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Baselinestudie, Triangulation	
Gesamtbeurteilung des Bewertungskriteriums			4

Maximal 5 Punkte: rot: 0 – 1 Punkte

gelb: 2 – 3 Punkte

grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Nein	Teilweise	Zu mild
Effektivität	Nein	Teilweise	Zu mild
Impact	Ja	Nein	Zu mild
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Nein	Teilweise	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			6

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte

gelb: 7 – 12 Punkte

grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	6

Maximal 4 Punkte: rot: 0 – 2 Punkt

gelb: 3 – 6 Punkte

grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	1	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien	1	
4. Empfehlungen	2	
Gesamtbewertung der Evaluierung	10	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Nein</i>	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Ja</i>	<i>Friedenspädagogik</i>
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>
		X				
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>B</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>	
			X			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X		X	

6.11.21 Analyseraster Meta-Evaluierung PFK überregional (Sektorvorhaben)

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Nein	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		6

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	4

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja		
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			3
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Ja	Ja	Angemessen
Impact	Ja	Teilweise	Angemessen
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Ja	Angemessen
Gesamtbeurteilung des Bewertungskriteriums			14

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch nur am Rande thematisiert)
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	1	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien	3	
4. Empfehlungen	3	
Gesamtbewertung der Evaluierung	15	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung						<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.						<i>Ja</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.						-	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.						-	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.						-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein-amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost-europa</i>	<i>Zentralasien</i>	
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>		
					<i>X</i>		
Art der Evaluierung			<i>Zwischen-evaluierung</i>	<i>Schluss-evaluierung</i>	<i>Ex-post Evaluierung</i>		
			<i>X</i>				
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>		
					<i>X</i>		
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>		
					<i>X</i>		

6.11.22 Analyseraster Meta-Evaluierung PFK Zentralasien

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Ja	Ja
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Nein	Nein
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Nein	Nein
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Gesamtbeurteilung des Bewertungskriteriums		6

Maximal 13 Punkte: rot: 0 – 6 Punkte gelb: 7 – 10 Punkte grün: 11 – 13 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Teilweise
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Nein
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Nein
Gesamtbeurteilung des Bewertungskriteriums	1

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	3
Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte	

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	u.a. diverse Länderberichte, M&E Berichte	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Nein		
Vergleichsgruppen wurden berücksichtigt.	Nein		Keine KG vorhanden
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			2
Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte			

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	Ja	Angemessen
Effektivität	Nein	Teilweise	Angemessen
Impact	Ja	Teilweise	Angemessen
Effizienz	Ja	Ja	Angemessen
Nachhaltigkeit	Ja	Teilweise	Zu mild
Gesamtbeurteilung des Bewertungskriteriums			11

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch nur wenig thematisiert)
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	7

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium	Bewertung	
1. Berücksichtigung der Wirkungsebenen	1	
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	1
	2.2 Angemessenheit von Design und Datenerhebung	3
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien	2	
4. Empfehlungen	3	
Gesamtbewertung der Evaluierung	12	

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 - 11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					<i>ja / nein</i>	<i>Welche?</i>
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					<i>Nein</i>	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					<i>Nein</i>	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					<i>Ja</i>	<i>Grundbildung Berufsbildung</i>
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	<i>Subsahara Afrika</i>	<i>Asien / Pazifik</i>	<i>Latein- amerika</i>	<i>Naher Osten / Nordafrika</i>	<i>Südost- europa</i>	<i>Zentralasien</i>
						X
Projekttyp			<i>Bilateral/Multilateral</i>	<i>Bilateral/Regional</i>	<i>Sektorvorhaben</i>	
			<i>B</i>	<i>R</i>		
Art der Evaluierung			<i>Zwischen- evaluierung</i>	<i>Schluss- evaluierung</i>	<i>Ex-post Evaluierung</i>	
			X			
Zeitraum der Durchführung			<i>Vor der Fusion</i>		<i>Nach der Fusion</i>	
			X		X	
Durchführungsorganisation			<i>GTZ</i>	<i>InWEnt</i>	<i>Fusionierte GIZ</i>	
			X		X	

6.11.23 Analyseraster Meta-Evaluierung DE Kosovo

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten	Nein	Nein
Leistungen (Outputs)	Ja	Ja
Nutzung der Leistungen (intermediate Outcome)	Ja	Ja
direkte Wirkungen (Outcome)	Ja	Ja
Impact	Ja	Ja
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	Nein
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	Nein
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	Nein
Gesamtbeurteilung des Bewertungskriteriums		8

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	4

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	Berichte von internationalen Organisationen, staatliche Dokumente etc.	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Interviews mit relevanten Stakeholdern	
Vergleichsgruppen wurden berücksichtigt.	Ja	ZG/KG Vergleich	
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Ja	Baselinestudie	
Gesamtbeurteilung des Bewertungskriteriums			5

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	-	Keine Notengebung vorhanden
Effektivität	Ja	-	Keine Notengebung vorhanden
Impact	Ja	-	Keine Notengebung vorhanden
Effizienz	Ja	-	Keine Notengebung vorhanden
Nachhaltigkeit	Ja	-	Keine Notengebung vorhanden
Gesamtbeurteilung des Bewertungskriteriums			5

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch nur marginal thematisiert)
Die Empfehlungen sind spezifisch.	Teilweise
Die Empfehlungen sind klar adressiert.	Teilweise
Gesamtbeurteilung des Bewertungskriteriums	6

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		2
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	3
3. Verwendung der OECD-DAC Kriterien		1
4. Empfehlungen		2
Gesamtbewertung der Evaluierung		13

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Nein	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Nein	
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					Ja	Jugendförderung
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
					X	
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	B		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
				X		
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
			X		X	

6.11.24 Analyseraster Meta-Evaluierung DE Zentralamerika

Bewertungskriterien

Berücksichtigung der Wirkungsebenen (M1, M2, M3)

Wirkungsebene	Ist die Wirkungsebene in der Evaluierung berücksichtigt? (ja = 1 Punkt)	Ist die Wirkungsebene in der Evaluierung richtig berücksichtigt? (ja = 1 Punkt)
Aktivitäten		
Leistungen (Outputs)		
Nutzung der Leistungen (intermediate Outcome)		
direkte Wirkungen (Outcome)		
Impact		
Wurde in der Evaluierung berücksichtigt, ob die Indikatoren nach SMART-Kriterien formuliert wurden?	teilweise = 1 Punkt, ja = 2 Punkte	
Die Wirkungshypothesen werden in der Evaluierung formuliert	ja = 1 Punkt	
Wenn ja: Die formulierten Wirkungshypothesen sind schlüssig	ja = 1 Punkt	
Gesamtbeurteilung des Bewertungskriteriums		<i>Nicht bewertbar</i>

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Methodisches Vorgehen (M4, M5, M6, M7, M8)

2.1 Transparenz im methodischen Vorgehen (M6)	
Die Arbeitsschritte im methodischen Vorgehen werden transparent beschrieben. (ja = 2 Punkte)	Ja
Die Daten und Angaben sind vollständig vorhanden. (ja = 1 Punkt)	Ja
Die Methodenwahl wird in der Evaluierung begründet. (ja = 2 Punkte)	Ja
Gesamtbeurteilung des Bewertungskriteriums	5

Maximal 5 Punkte: rot: 0 – 1 Punkt gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

2.2 Angemessenheit von Design und Datenerhebung (M7)	ja = 1 Punkt
Design und Datenerhebung berücksichtigen die Interventionsebenen (Mikro-, Meso-, Makroebene) des evaluierten Vorhabens.	Nein
Design und Datenerhebung berücksichtigen die Projektphasen (eine Projektphase, mehrere Projektphasen) des evaluierten Vorhabens.	Ja
Design und Datenerhebung berücksichtigen die regionalen Dimensionen (scope) des evaluierten Vorhabens.	Ja
Gesamtbeurteilung des Bewertungskriteriums	2

Maximal 5 Punkte: rot: 0 Punkte gelb: 1 – 2 Punkte grün: 3 Punkte

2.3 Adäquate Datenbasis (M4, M5, M8)	ja = 1 Punkt	Welche	Kommentare
Die Evaluierungsergebnisse basieren auf Dokumentenstudium	Ja		
Die Evaluierungsergebnisse basieren auf empirischen Sekundärdaten.	Ja	u.a. Evaluationen, Broschüren, DVDs	
Die Evaluierungsergebnisse basieren auf empirischen Primärdaten.	Ja	Leitfadengestützte Interviews,	
Vergleichsgruppen wurden berücksichtigt.	Nein		
Vorher-Nachher-Vergleiche wurden berücksichtigt.	Nein		
Gesamtbeurteilung des Bewertungskriteriums			3

Maximal 5 Punkte: rot: 0 – 1 Punkte gelb: 2 – 3 Punkte grün: 4 – 5 Punkte

Verwendung der OECD-DAC Kriterien (M9)

Kriterium*	Die Evaluierungskriterien wurden gemäß Vorgaben richtig angewendet. (ja = 1 Punkte)	Die Argumentation, welche zum Gesamturteil führt, ist nachvollziehbar. (teilweise = 1 Punkt, ja = 2 Punkt)	Beurteilung der Benotung (zu mild; angemessen, zu streng)
Relevanz	Ja	-	Keine Notengebung vorhanden
Effektivität	Ja	-	Keine Notengebung vorhanden
Impact	Ja	-	Keine Notengebung vorhanden
Effizienz	Ja	-	Keine Notengebung vorhanden
Nachhaltigkeit	Ja	-	Keine Notengebung vorhanden
Gesamtbeurteilung des Bewertungskriteriums			5

* Die Ergebnisse pro Kriterium über alle Evaluierungen werden in der Gesamtauswertung dargestellt.

Maximal 15 Punkte: rot: 0 – 6 Punkte gelb: 7 – 12 Punkte grün: 13 – 15 Punkte

Die Frage nach der Einheitlichkeit der Notengebung wird anhand der Gesamtbeurteilung dieses Kriteriums beantwortet. (M9)

Empfehlungen

Empfehlungen	ja = 2 Punkte teilweise = 1 Punkt
Die in der Evaluierung ausgesprochenen Empfehlungen sind aufgrund der Evaluierungsergebnisse nachvollziehbar.	Ja
Die in der Evaluierung ausgesprochenen Empfehlungen berücksichtigen den aktuellen Wissensstand in der Bildung.	Ja (jedoch kaum thematisiert)
Die Empfehlungen sind spezifisch.	Ja
Die Empfehlungen sind klar adressiert.	Ja
Gesamtbeurteilung des Bewertungskriteriums	8

Maximal 4 Punkte: rot: 0 – 2 Punkt gelb: 3 – 6 Punkte grün: 7 – 8 Punkte

Gesamtbewertung der Einzelevaluierung

Kriterium		Bewertung
1. Berücksichtigung der Wirkungsebenen		1
2. Methodisches Vorgehen	2.1 Transparenz im methodischen Vorgehen	3
	2.2 Angemessenheit von Design und Datenerhebung	2
	2.3 Adäquate Datenbasis	2
3. Verwendung der OECD-DAC Kriterien		1
4. Empfehlungen		3
Gesamtbewertung der Evaluierung		12

Skala für die Beurteilung der einzelnen Kriterien: rot = 1 Punkt, gelb = 2 Punkte, grün = 3 Punkte

Skala für die Gesamtbewertung der Evaluierung

Rot: Durchschnittsbewertung 6 -11

Gelb: Durchschnittsbewertung 12 – 15

Grün: Durchschnittsbewertung 16 – 18

Aspekte zur Vergleichbarkeit

Zuordnungsfragen Sektor Bildung					ja / nein	Welche?
Das evaluierte Vorhaben nimmt den gesamten Sektor Bildung in den Blick.					Nein	
Das evaluierte Vorhaben nimmt Teilaspekte des Sektors Bildung in den Blick.					Nein	
Im evaluierten Vorhaben ist Bildung eine Subkomponente.					Ja	Friedenspädagogik
Das evaluierte Vorhaben hat erwartete Wirkungen auf den Sektor Bildung.					-	
Region	Subsahara Afrika	Asien / Pazifik	Latein-amerika	Naher Osten / Nordafrika	Südost-europa	Zentralasien
			X			
Projekttyp			Bilateral/Multilateral	Bilateral/Regional	Sektorvorhaben	
			B	R		
Art der Evaluierung			Zwischen-evaluierung	Schluss-evaluierung	Ex-post Evaluierung	
				X		
Zeitraum der Durchführung			Vor der Fusion		Nach der Fusion	
			X		X	
Durchführungsorganisation			GTZ	InWEnt	Fusionierte GIZ	
				X	X	

6.12 Literatur

- Atteslander, P. (2010). *Methoden der empirischen Sozialforschung* (13. Aufl.). Berlin: Schmidt.
- Bergmann, H. (2012). *Desk Study: Die Evaluierung von Grundbildungsvorhaben – ein Orientierungsrahmen*. Eschborn.
- Beywl, W. (2008). *Standards für Evaluation*. (DeGEval - Gesellschaft für Evaluation e.V., Hrsg.). Köln: DeGEval - Gesellschaft für Evaluation.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Cross, M., Mungadi, R. & Rouhani, S. (2002). From policy to practice: Curriculum reform in South African education. *Comparative Education*, 38(2), 171–187.
- Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York: Routledge.
- Houlmann, L. (2014). What is NORRAG? *NORRAG (Network for International Policies and Cooperation in Education and Training)*. Abgerufen von <http://norrags.wordpress.com/about/>
- Huber, S. G. (Hrsg.). (in Vorbereitung). *International Handbook of Evaluation in Education*. Dordrecht: Springer.
- Johnson, D. (2006). Comparing the trajectories of educational change and policy transfer in developing countries. *Oxford Review of Education*, 32(5), 679–696.
- Klieme, E. (2007). *Zur Entwicklung nationaler Bildungsstandards : eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- Komarek, K. (2006). „Bildung für alle“ durch muttersprachlichen Unterricht. *eins Entwicklungspolitik*, (18/19), 30–33.
- Lange, S. (2014). Learner orientation through professional development of teachers? Empirical results from cascade training in Anglophone Cameroon. *Compare: A Journal of Comparative and International Education*, 44(4), 587–612. doi:10.1080/03057925.2013.841027

- Meuser, M. & Nagel, U. (1991). ExpertInneninterviews–vielfach erprobt, wenig bedacht. In D. Garz & K. Kraimer (Hrsg.), *Qualitativ-empirische Sozialforschung: Konzepte, Methoden, Analysen* (S. 441–471). Opladen: Westdeutscher Verlag.
- Meuser, M. & Nagel, U. (2005). Vom Nutzen der Expertise: Experteninterviews in der Sozialberichterstattung. In A. Bogner & W. Menz (Hrsg.), *Das Experteninterview: Theorie, Methode, Anwendung* (S. 257–272). VS Verlag für Sozialwissenschaften.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Thousand Oaks, CA: SAGE Publications.
- Sanders, J. R. & Beywl, W. (2006). *Handbuch der Evaluationsstandards: die Standards des „Joint Committee on Standards for Educational Evaluation“* (3., erw. und aktualisierte Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schnell, R., Hill, P. B. & Esser, E. (2008). *Methoden der empirischen Sozialforschung* (8. Aufl.). München: Oldenbourg Wissenschaftsverlag.
- Scriven, M. (1991). *Evaluation Thesaurus* (4. Aufl.). Newbury Park, CA: SAGE.
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004, Oktober 15). Vereinbarung über Bildungsstandards für den Hauptschulabschluss (Jahrgangsstufe 9). Abgerufen von http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Haupt.pdf
- Steiner-Khamsi, G. (2006). The economics of policy borrowing and lending: A study of late adopters. *Oxford Review of Education*, 32(5), 665–678.