

WISSEN WAS WIRKT

# Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben

Querschnittsauswertung - Hauptbericht  
Im Auftrag der GIZ durchgeführt von externen Evaluierenden

## Impressum

Als Bundesunternehmen unterstützt die GIZ die deutsche Bundesregierung bei der Erreichung ihrer Ziele in der Internationalen Zusammenarbeit für nachhaltige Entwicklung.

Die Stabsstelle Evaluierung der GIZ untersteht organisatorisch direkt dem Vorstand. Diese Organisationsstruktur stellt ihre Unabhängigkeit vom operativen Geschäft sicher. Die Stabsstelle Evaluierung generiert evidenzbasierte Ergebnisse und Empfehlungen. Sie sorgt für Transparenz der Erkenntnisse aus Evaluierungen und fördert deren Nutzung.

Diese Evaluierung wurde im Auftrag der Stabsstelle Evaluierung von externen Evaluierenden durchgeführt und der Evaluierungsbericht von externen Evaluierenden verfasst. Er gibt ausschließlich deren Meinung und Wertung wieder.

### Evaluierende :

Lennart Raetzell und Alexandra Hoppe, Syspons GmbH

### Autor\*innen:

Lennart Raetzell und Alexandra Hoppe, Syspons GmbH

### Consulting:

Syspons GmbH  
Prinzenstr. 84 Aufgang 1  
10969 Berlin / Germany  
T: +49 30 698 158 00  
E: info@syspons.com  
I: www.syspons.com



### Konzeption, Koordination und Management

Claudia Kornahrens, GIZ, Gruppenleiterin  
Benjamin Bräuer, GIZ, Spezialist Evaluierung  
Lucas Jacobs, GIZ, Spezialist Evaluierung  
Gruppe Zentrale Projektevaluierung  
Stabsstelle Evaluierung der GIZ

### Verantwortlich:

Albert Engel, GIZ, Leiter  
Stabsstelle Evaluierung der GIZ

### Herausgeberin:

Deutsche Gesellschaft für  
Internationale Zusammenarbeit (GIZ) GmbH

### Sitz der Gesellschaft

Bonn und Eschborn

Friedrich-Ebert-Allee 32 + 36  
53113 Bonn, Deutschland  
T +49 228 4460-0  
F +49 228 4460 - 1766

E [evaluierung@giz.de](mailto:evaluierung@giz.de)  
I [www.giz.de/wissenwaswirkt](http://www.giz.de/wissenwaswirkt)  
[www.youtube.com/user/GIZonlineTV](https://www.youtube.com/user/GIZonlineTV)  
[https://twitter.com/giz\\_gmbh](https://twitter.com/giz_gmbh)

### Design/Layout etc.:

DITHO Design GmbH, Köln  
now [nau], kommunikative & visuelle gestaltung,  
Frankfurt

### Vertrieb:

GIZ, Bonn

Bonn, November 2022

## Zusammenfassung

Die Querschnittsauswertung „Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben“ wurde auf der Basis einer Zufallsstichprobe von 50 Zentralen Projektevaluierungen und vertiefender Interviews durchgeführt. Sie zielte darauf ab, Erfolgs- und Misserfolgskriterien bei der Verzahnung dieser drei Prozesse zu identifizieren.

Ein wesentlicher Anspruch an die Planung der Vorhaben ist, dass diese dazu beiträgt, ein effektives Wirkungsorientiertes Monitoring (WoM) der Vorhaben zu etablieren. Bei der Verzahnung von **Planung und Monitoring** wurden die Aspekte Wirkungsmodelle, Indikatoren und Datenquellen betrachtet. Folgende Erfolgs- und Misserfolgskriterien wurden benannt:

- Zusammenfassend ist in der qualitativen Auswertung zu erkennen, dass sich die bisherigen Vorgaben/Handreichungen sowie das Wirkungsmodell als Erfolgsfaktoren für eine Synchronisierung zwischen Planung und Monitoring herausgestellt haben. Im Gegensatz dazu haben sich in der Auswertung die für das Monitoring bereitgestellten, zu geringen, Ressourcen sowie das lückenhafte Verständnis in Bezug auf konflikt-sensibles Monitoring als Misserfolgskriterien für eine Synchronisierung zwischen Planung und Monitoring herausgestellt.

Das Monitoring der Vorhaben soll nicht nur zur Steuerung des Vorhabens, sondern auch zu dessen Evaluierung beitragen. Die Erkenntnisse zur Synchronisierung zwischen **Monitoring und Evaluierung** der ausgewählten Vorhaben sind dabei entsprechend dem Monitoring der Vorhaben, dem Monitoring der Wirkungshypothesen, den Erkenntnissen zu Beiträgen der Vorhaben auf der Impact-/Programm-Ebene und der Bewertung des WoM der Vorhaben sortiert. Folgende Erfolgs- und Misserfolgskriterien wurden benannt:

- Die bisherigen Vorgaben/Handreichungen sowie das gemeinsame Verständnis im Projektteam, das Monitoring relevant für die Projektsteuerung ist, gelten als Erfolgsfaktoren für eine Synchronisierung zwischen Monitoring und Evaluierung. In der Auswertung haben sich erneut die für das Monitoring bereitgestellten Ressourcen sowie das lückenhafte Verständnis in Bezug auf konflikt-sensibles Monitoring als Misserfolgskriterien für eine Synchronisierung zwischen Monitoring und Evaluierung herausgestellt.

Eine Zielsetzung hinter den Evaluierungen der STS Evaluierung ist, dass diese die Planung von Vorhaben der GIZ unterstützen. Die Querschnittsauswertung hat daher bezüglich **Evaluierung und Planung** die Frage untersucht, welche Potenziale (ZPE-)Evaluierungsergebnisse auf den verschiedenen Ebenen innerhalb der GIZ für die Planung von Neu- und/oder Folgevorhaben sowie für das Lernen (im Land, Sektor und/oder Organisationales Lernen) bieten, und wie sie genutzt werden. Dabei wurden zuerst die Erkenntnisse zur Implementierung der Wirkungshypothesen betrachtet, gefolgt von Erkenntnissen zur Verzahnung mit Vorgängervorhaben und zur Nutzung von Evidenzen für Folgevorhaben.

- In der Auswertung hat sich das rechtzeitige Bereitstellen von relevanter Information als Erfolgsfaktor für eine Synchronisierung zwischen Evaluierung und Planung herausgestellt. In der Auswertung hat sich der zeitliche Rahmen der ZPE und Planung von (Folge-)Vorhaben sowie eine fehlende Vergemeinschaftung von Evaluierungsergebnissen als Misserfolgswfaktor für eine Synchronisierung zwischen Evaluierung und Planung herausgestellt.

Die aus der Analyse und den Schlussfolgerungen abgeleiteten Handlungsempfehlungen beziehen sich insbesondere auf eine stärkere Berücksichtigung eines konfliktensensiblen Monitorings, eine ausreichende Bereitstellung von Personalressourcen für das Monitoring sowie auf eine stärkere Verankerung der Nutzung von Evidenzen im Planungs- und Prüfprozess, unterstützt durch eine Aufbereitung und Vergemeinschaftung der Evaluierungsergebnisse zum Beispiel in Form eines Dashboards.

# Inhalt

<b>Zusammenfassung .....</b>	<b>3</b>
<b>Abbildungsverzeichnis.....</b>	<b>4</b>
<b>Abkürzungsverzeichnis.....</b>	<b>6</b>
<b>1 Einleitung.....</b>	<b>7</b>
<b>2 Methodisches Vorgehen .....</b>	<b>8</b>
2.1 Untersuchungsgegenstand .....	8
2.2 Datengrundlage und Evidenzstärke .....	9
2.3 Auswertungsprozess .....	10
<b>3 Ergebnisse der Querschnittsauswertung.....</b>	<b>12</b>
3.1 Planung - Monitoring .....	12
Wirkungsmodelle .....	12
Indikatoren.....	15
Datenquellen .....	17
Erfolgs- und Misserfolgskfaktoren .....	19
3.2 Monitoring - Evaluierung .....	20
Monitoring der Vorhaben .....	20
Monitoring der Wirkungshypothesen .....	26
Beiträge auf der Impact-/Programm-Ebene .....	27
Bewertung des WoM .....	28
Evaluability Assessment.....	29
Erfolgs- und Misserfolgskfaktoren.....	30
3.3 Evaluierung - Planung .....	30
Implementierung der Wirkungshypothesen .....	30
Vorgängervorhaben.....	32
Folgevorhaben.....	33
<b>4 Schlussfolgerungen und Empfehlungen .....</b>	<b>35</b>
<b>Literatur .....</b>	<b>40</b>
<b>Anhang.....</b>	<b>41</b>
Anhang 1: Analyseraster .....	41
Anhang 2: Erhebungsmaske .....	41

## Abbildungsverzeichnis

Abbildung 1 Synchronisierung von Planung, Monitoring und Evaluierung der Vorhaben der GIZ .....	7
Abbildung 2: Wirkungsmodell der betrachteten Vorhaben.....	13
Abbildung 3: Nutzung der Wirkungsmodelle der betrachteten Vorhaben in den ZPE .....	13
Abbildung 4: Systemgrenze der betrachteten Vorhaben .....	14
Abbildung 5: Untersuchung der Elemente der Wirkungsmodelle der betrachteten Vorhaben in der ZPE.....	14
Abbildung 6: SMART-Bewertung der Indikatoren der betrachteten Vorhaben in der ZPE .....	16
Abbildung 7: SMART-Anpassung der Indikatoren der betrachteten Vorhaben in der ZPE .....	16
Abbildung 8: Anzahl der SMART-angepassten Indikatoren der betrachteten Vorhaben in der ZPE.....	17
Abbildung 9: SMART-Kategorie, die in der ZPE der betrachteten Vorhaben angepasst wurde.....	17
Abbildung 10: Anzahl der vorliegenden Baselinewerte für die Indikatoren der untersuchten Vorhaben .....	18
Abbildung 11: Nachvollziehbarkeit der Baselinewerte für die Indikatoren der untersuchten Vorhaben.....	18
Abbildung 12: Art der Datenerhebung der Baselinewerte für die Indikatoren der untersuchten Vorhaben .....	18
Abbildung 13: Nachvollziehbarkeit der Zielwerte der Indikatoren der untersuchten Vorhaben .....	19
Abbildung 14: Art der Datenerhebung im laufenden Monitoring der untersuchten Vorhaben .....	21
Abbildung 15: Regelmäßigkeit der selbstständig durchgeführten Monitoringerhebungen der untersuchten Vorhaben .....	21
Abbildung 16: Regelmäßigkeit der selbstständig durchgeführten Monitoringmeetings in den untersuchten Vorhaben .....	22
Abbildung 17: Nutzung der Monitoringdaten zur Steuerung der untersuchten Vorhaben .....	23
Abbildung 18: Erhebung von nicht-intendierten Wirkungen im Monitoring der untersuchten Vorhaben.....	24
Abbildung 19: Abbildung von Querschnittsthemen in den Indikatoren der untersuchten Vorhaben.....	25
Abbildung 20: Abbildung der Wirkungshypothesen in den darunterliegenden Indikatoren der untersuchten Vorhaben .....	26
Abbildung 21: Berücksichtigung der Programm-Ebene in den betrachteten Vorhaben.....	27
Abbildung 22: Nutzung der Monitoringdaten in den ZPE .....	28
Abbildung 23: Nutzung der Monitoringdaten in den ZPE zur Validierung der Wirkungshypothesen (Outcome-Ebene) .....	28
Abbildung 24: Nutzung der Monitoringdaten in den ZPE zur Validierung der Wirkungshypothesen (Impact-Ebene) .....	29
Abbildung 25: <i>Evaluability Assessment</i> (Ausschnitt) (Quelle: GIZ). .....	29
Abbildung 26: Umsetzung der Wirkungshypothesen (Outcome-Ebene) der untersuchten Vorhaben.....	31
Abbildung 27: Änderungen an Wirkungshypothesen (Outcome-Ebene) der untersuchten Vorhaben in der ZPE .....	31
Abbildung 28: Umsetzung der Wirkungshypothesen (Impact-Ebene) der untersuchten Vorhaben.....	32

Abbildung 29: Änderungen an Wirkungshypothesen (Impact-Ebene) der untersuchten Vorhaben in der ZPE .	32
Abbildung 30: Nutzung der Erfahrungen der Vorgängervorhaben.....	33
Kasten 1: Beispiele für gute Praktiken: Art der Datenerhebung im laufenden Monitoring .....	21
Kasten 2: Beispiele für gute Praktiken: Regelmäßig selbstständig durchgeführte Monitoringerhebungen .....	22
Kasten 3: Beispiele für gute Praktiken: Regelmäßige selbstständig durchgeführte Monitoringmeetings .....	23
Kasten 4: Beispiele für gute Praktiken: Zur Steuerung genutzte Monitoringdaten .....	24
Kasten 5: Beispiel: Erhebung nicht-intendierter Wirkungen im Monitoring .....	25
Kasten 6: Beispiele: Abbildung von Querschnittsthemen in den Indikatoren der untersuchten Vorhaben .....	26
Kasten 7: Beispiel: Indikator, der Teil einer Wirkungshypothese abbildet .....	27
Kasten 8: Beispiel für gute Praktiken: Nutzung der Erfahrungen der Vorgängervorhaben.....	33
Kasten 9: Ausgewählte Rückmeldungen der AV-(V-)Befragung: Nutzung der ZPE-Ergebnisse im Folgevorhaben .....	33
Kasten 10: Ausgewählte Rückmeldungen der AV-(V)Befragung: Nutzung des eigenen Erfahrungsschatzes...	34

## Abkürzungsverzeichnis

AV	Auftragsverantwortliche*r
AV-V	Auftragsverantwortliche*r in der Vorbereitung
BMZ	Bundesministerium für Wirtschaftliche Zusammenarbeit und Entwicklung
DAC	<i>Development Assistance Committee</i> Entwicklungsausschuss
FMB	Fach- und Methodenbereich
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH
OECD	Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
PME	Planung, Monitoring und Evaluierung
PN	Projektnummer
SMART	spezifisch, messbar, erreichbar (achievable), relevant, zeitgebunden (time-bound)
STS	Stabsstelle Evaluierung der GIZ
ToC	Theory of Change
WoM	Wirkungsorientiertes Monitoring
ZPE	Zentrale Projektevaluierung

# 1 Einleitung

## Hintergrund und Ziele der Querschnittsauswertung

Die Syspons GmbH wurde von der Deutschen Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ) damit beauftragt, die **Querschnittsauswertung (QSA)** „Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben“ (kurz: QSA PME) durchzuführen. Die QSA sollte dazu beitragen, Verbesserungspotenziale und Evidenzen bezüglich der Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben in der GIZ zu identifizieren.

In Querschnittsauswertungen werden Erkenntnisse aus Evaluierungen der GIZ analytisch aufbereitet und zusammenfassend dargestellt, um diese für das unternehmensweite Lernen nutzbar zu machen. Die Querschnittsauswertung *QSA PME* hatte das Ziel, vertiefte Erkenntnisse aus dem Evaluierungsinstrument der **Zentralen Projektevaluierungen (ZPE)** zu gewinnen. Mit den zentralen Projektevaluierungen erfasst die GIZ seit 2017/2018 regelmäßig und systematisch die Wirkung, Wirtschaftlichkeit und Nachhaltigkeit von Vorhaben, die sie zusammen mit Partnerorganisationen im Auftrag des Bundesministeriums für Wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) umsetzt. Gegenstand der ZPE im BMZ-Geschäft sind Projekte, die die GIZ im Auftrag des BMZ durchführt, um entweder Veränderungsprojekte der Partnerorganisationen zu unterstützen oder Sektor- und Globalfragen zu bearbeiten. Die Bewertung erfolgt entlang der sechs OECD/DAC-Kriterien Relevanz, Kohärenz (seit Beginn 2021), Effektivität, Effizienz, Impact und Nachhaltigkeit.

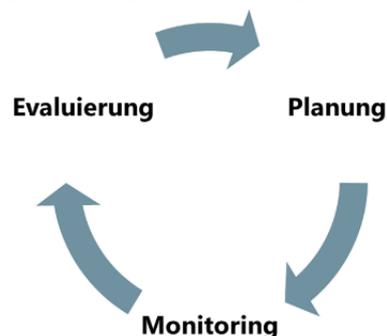
Grundsätzlich sollen Evaluierungen der vom BMZ beauftragten Projekte in der GIZ drei Grundfunktionen dienen: Transparenz und Rechenschaftslegung, Unterstützung evidenzbasierter Entscheidungen sowie einen Beitrag zu Wissensmanagement und organisationalem Lernen zu leisten. Die Planung, Durchführung und Nutzung von Evaluierungen gestaltet die GIZ daher so, dass der Beitrag des Evaluierungsprozesses und der Evaluierungsergebnisse zu diesen drei Grundfunktionen optimiert wird.

Vor diesem Hintergrund sollten durch die QSA konkret die folgenden **Ziele** erreicht werden:

- Besseres (empirisches) Verständnis bzgl. der Verzahnung von Planung, Monitoring und Evaluierung;
- Identifizierung von Herausforderungen;
- Ausarbeitung konkreter Empfehlungen zur Verbesserung.

Inhaltlich ist sie dabei entlang der Dimensionen Planung-Monitoring, Monitoring-Evaluierung, und Evaluierung-Planung strukturiert (siehe Abbildung 1).

Abbildung 1 Synchronisierung von Planung, Monitoring und Evaluierung der Vorhaben der GIZ



Quelle: Syspons 2021, basierend auf GIZ 2021

Der Bericht ist wie folgt aufgebaut: In Kapitel 2 wird das methodische Vorgehen der Querschnittsauswertung dargestellt. In diesem Kapitel wird zunächst der Untersuchungsgegenstand beschrieben. Dann werden die Datengrundlage und die Datenqualität beziehungsweise Evidenzstärke diskutiert und eingeordnet. Anschließend wird der Auswertungsprozess erläutert. In Kapitel 3 werden die Ergebnisse der Auswertung und Analyse vorgestellt und diskutiert. In Kapitel 4 werden Schlussfolgerungen und Empfehlungen dargestellt. Kapitel 5 enthält den Anhang.

## 2 Methodisches Vorgehen

In diesem Kapitel wird das methodische Vorgehen der Querschnittsauswertung: „Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben“ vorgestellt. Im folgenden Abschnitt wird zunächst der Untersuchungsgegenstand beschrieben. Anschließend werden die Datengrundlage und die Datenqualität in Bezug auf die Auswertung eingeordnet. Die einzelnen Schritte der Auswertung werden im letzten Abschnitt über den Auswertungsprozess dargestellt.

### 2.1 Untersuchungsgegenstand

Der **zentrale Untersuchungsgegenstand** zur Auswertung und Analyse der Synchronisierung von Planung, Monitoring und Evaluierung von Vorhaben waren **50 Zentrale Projektevaluierungen**.

Insgesamt waren bis zum Beginn der QSA PME 122 Vorhaben evaluiert worden. Für die Auswahl der durch die ZPE zu evaluierenden Vorhaben arbeitet die GIZ mit einer mit dem BMZ abgestimmten **repräsentativen Stichprobe**. Eine jährliche (geschichtete) Zufallsstichprobe von etwa 40 Prozent der Grundgesamtheit – alle BMZ-finanzierten Vorhaben mit einem Auftragswert von mindestens drei Millionen Euro, die im Folgejahr enden – ermöglicht dabei statistisch repräsentative Aussagen über die Erfolgsquote aller in einem **Zeitraum von zwei Jahren** abgeschlossenen Vorhaben. Ergänzt wird das ZPE-Portfolio durch Evaluierungen von Vorhaben, die aufgrund einer besonderen strategischen, politischen oder auch methodischen Relevanz vorgeschlagen werden.

Bei der Auswahl der Evaluierungsberichte für die Betrachtung in Bezug auf die Synchronisierung von Planung, Monitoring und Evaluierung handelte es sich ebenfalls um eine geschichtete **Zufallsstichprobe**. Bei der Ziehung der geschichteten Zufallsstichprobe wurde darauf geachtet, dass eine Verteilung der evaluierten Vorhaben nach Region, nach Berichtsjahr<sup>1</sup> sowie eine Abbildung fragiler Kontexte<sup>2</sup> gegeben war.

---

<sup>1</sup> In der Stichprobe waren ZPE-Berichte von 2018 bis 2021 enthalten.

<sup>2</sup> Die Zuteilung zu fragilen Kontexten wurde dabei nach dem Fragile State Index vorgenommen. Hierbei wurde geprüft, ob die Vorhaben Aktivitäten in Ländern durchführen, die laut Fragile State Index als fragil gelten. Nach dieser Zuteilung führten 13 der 50 evaluierten Vorhaben (26 %) der Stichprobe Aktivitäten in fragilen Kontexten durch.

## 2.2 Datengrundlage und Evidenzstärke

Die **zentrale Datengrundlage** der Untersuchung bildet eine Stichprobe von 50 der 122 zu Beginn der QSA PME abgeschlossenen Zentralen Projektevaluierungen von durch das BMZ beauftragten Vorhaben (siehe Kapitel 2.1). Die Evaluierungen wurden von unabhängigen Evaluatord\*innen durchgeführt, welche von der Stabsstelle (STS) Evaluierung der GIZ beauftragt worden waren. Die Stichprobe wurde dabei als geschichtete Zufallsstichprobe ausgewählt, um die Robustheit der Ergebnisse sicherzustellen. Zusätzlich wurde die dazugehörige Qualitätsbewertung der Evaluierungen durch die STS Evaluierung betrachtet, um zu erheben, inwiefern ein gutes Wirkungsorientiertes Monitoring (WoM) mit der Qualität der Evaluierung korreliert. Ergänzend wurden elf weitere Dokumente wie GIZ-interne Studien, die Checkliste *Evaluability Assessment*, die Prozessbeschreibung und Checkliste für die Prüfung, sowie ausgewählte Angebote beziehungsweise Folgeangebote und Monitoringsysteme ausgewertet. Zusätzlich zur Auswertung der Dokumente hat Syspons einerseits explorative, vertiefende Interviews mit Vertreter\*innen der STS Unternehmensentwicklung sowie des Fach- und Methodenbereichs (FMB) der GIZ und andererseits im Rahmen sogenannter Deep Dives Interviews mit den Auftragsverantwortlichen (AVs) beziehungsweise Auftragsverantwortlichen der Vorbereitung (AV-Vs) sowie Monitoringbeauftragten von fünf ausgewählten Vorhaben durchgeführt. Außerdem hat die Stabsstelle Evaluierung eine informelle E-Mail-Anfrage an ausgewählte AVs beziehungsweise AV-Vs von Vorhaben mit Folgevorhaben gerichtet und darauf 15 Rückmeldungen erhalten, die in die Auswertung eingeflossen sind.

Zur **Evidenzstärke** der Analyse und Bewertung der Synchronisierung von Planung, Monitoring und Evaluierung der Vorhaben ist anzumerken, dass diese besonders von der Aussagekraft der Evaluierungsberichte, der ausgewählten Dokumente sowie der Interviews und/oder Rückmeldungen der AVs sowie der weiteren Projektteams abhing. Mit Blick auf die ZPE-Berichte können zum Beispiel die im Laufe der ZPE vorgenommenen Änderungen an Wirkungsmodellen der Vorhaben nur dann erhoben und ausgewertet werden, wenn in den Evaluierungen direkt oder zumindest indirekt darauf Bezug genommen wurde. Es wurden nur jene Dokumente im Rahmen der Dokumentenanalyse in die Auswertung einbezogen, die vorher durch die Stabsstelle Evaluierung und/oder im Rahmen der explorativen Interviews als relevant identifiziert worden waren. Bezüglich der Projektteams hängt die Aussagekraft der Interviews und/oder Rückmeldungen unmittelbar von den beteiligten Personen ab.

Die Datengrundlage für die QSA kann an dieser Stelle insgesamt als *gut* bezeichnet werden. So werden die ZPE sowohl von unabhängigen Evaluationsteams, das heißt von externen Evaluatord\*innen im Auftrag der Stabsstelle Evaluierung der GIZ, als auch nach gleichen Standards in Bezug auf Format, Inhalt und Qualität durchgeführt. Die Datengrundlage der Dokumentenanalyse, also die als relevant identifizierten Handreichungen und/oder Vorgaben, basierte auf Gesprächen mit Angehörigen verschiedener Abteilungen der GIZ (FMB, STS Unternehmensentwicklung, STS Evaluierung) und decken damit sämtliche Perspektiven beziehungsweise Abschnitte des Planungs- und Evaluierungsprozesses von Vorhaben ab. Bei den Projektteams wurde die Aussagekraft ihrer Rückmeldungen dadurch gestärkt, dass sie ausführlich über den Zweck der Untersuchung aufgeklärt wurden. So erhielten sie beispielsweise Gesprächsleitfäden, die der Vorbereitung und Orientierung während der Interviews dienten und in denen der Zweck der QSA sowie des Interviews dargelegt wurden.

## 2.3 Auswertungsprozess

### Vorgehen in der Vorbereitungsphase

Das Ziel der Vorbereitungsphase war es, mit der GIZ ein gemeinsames Verständnis der Zielsetzung und Umsetzung der QSA zu entwickeln. Dies bedeutete auf der Prozessebene, Erwartungen an den Nutzen der QSA zu klären und die Rollenverständnisse der Beteiligten abzustimmen. Auf der fachlich-methodischen Ebene bedeutete dies, einen Überblick über die Datenlage zu gewinnen und ein analyseleitendes Untersuchungssystem zu etablieren, das alle wesentlichen Aspekte und Fragestellungen abdeckt.

Zu diesem Zweck führte das Syspons-Team zunächst ein ausführliches Auftragsklärungsgespräch mit der Stabsstelle Evaluierung der GIZ als Auftraggeberin der Auswertung. Auf dieser Grundlage wurden das Analyseraster (siehe Anhang 2) sowie die *Concept Note* entwickelt.

Im Anschluss an die *Concept Note* wurden vertiefende Interviews mit Vertreter\*innen des Fach- und Methodenbereichs (FMB) sowie der STS Unternehmensentwicklung der GIZ durchgeführt. Diese Interviews dienten einerseits dem Zweck, weitere Beobachtungen und Erkenntnisinteressen zu erfragen, und das Analyseraster anzupassen. Andererseits wurde in ihnen nach Beispielen für gute Praktiken bei der Übertragung von Evaluierungsergebnissen in die Planung von Folgevorhaben gefragt.

### Vorgehen in der Analysephase

Ziel der Analysephase war es, eine quantitative und qualitative Datengrundlage für die QSA zu schaffen, die Untersuchungsfragen zu analysieren und zu beurteilen und die gesammelten Daten systematisch auszuwerten und zu analysieren.

Das Vorgehen in der Analysephase bestand aus vier methodischen Schritten: (1) qualitative Auswertung der ZPE-Berichte, (2) *Desk Research* unter anderem der GIZ-internen Prozessdokumente, (3) informelle E-Mail-Anfrage der STS Evaluierung an ausgewählte AVs beziehungsweise AV-Vs von Vorhaben mit Folgevorhaben, und (4) fünf Deep Dives, die die vertiefte Betrachtung von ausgewählten Vorhaben inkl. Durchführung von Interviews umfassten.

#### 1. Qualitative Auswertung der ZPE-Evaluierungsberichte

##### Ziehung der Stichprobe

In der Querschnittsauswertung wurden 50 (aus 122) zufällig ausgewählte ZPE-Berichte untersucht. Bei der Auswahl der Evaluierungsberichte handelte es sich um eine geschichtete **Zufallsstichprobe**. Bei der Ziehung der geschichteten Zufallsstichprobe wurde darauf geachtet, dass eine Verteilung der evaluierten Vorhaben nach Region, Berichtsjahr sowie eine Abbildung der fragilen Kontexte gegeben war.

##### Erstellung der Erhebungsinstrumente und Programmierung

Die ausgewählten 50 ZPE-Berichte wurden für die Auswertung zu möglichst gleichen Anteilen im Auswertungsteam verteilt. Mit einer **Erhebungsmaske** wurde dabei festgelegt, nach welchen konkreten Fragestellungen die ZPE-Berichte untersucht, verglichen und synthetisiert werden sollen. Die Erhebungsmaske baute auf dem Analyseraster auf und wurde in Abstimmung mit der STS Evaluierung entwickelt. Sie enthielt sowohl quantifizierte Fragen mit Kategorien (Ja/Nein/Teilweise/Keine Angaben) als auch Freitextfelder für die qualitativen Beobachtungen des Auswertungsteams. Nach der Fertigstellung der Auswertungsmaske wurde sie mithilfe des Programms *SurveyXact* für die computergestützte Auswertung der ZPE-Berichte programmiert (siehe nächster Abschnitt).

## Durchführung der qualitativen und quantitativen Auswertung der ZPE-Berichte

Im Rahmen der **computergestützten qualitativen und quantitativen Auswertung** hat das Auswertungsteam die ausgewählten ZPE-Berichte entlang der digitalisierten Erhebungsmaske betrachtet. Dabei wurde durch eine Schulung im Team zu Beginn der Auswertung ein gemeinsames Verständnis des Untersuchungsgegenstandes und der zu analysierenden Aspekte entwickelt und gestärkt. Während des Auswertungsprozesses bestand eine fortlaufende Kommunikation im Team. Bei Unklarheiten wurde der Dialog mit den anderen Teammitgliedern gesucht. Die für die Qualitätssicherung zuständige Person hat im Zweifelsfall Entscheidungen des Auswertungsteams geprüft.

Im Anschluss wurden die erhobenen Daten ausgewertet. Für die Auswertung wurden die gesammelten Daten sowohl quantitativ (Häufung von Beobachtungen in den Berichten) als auch qualitativ (Art der Beobachtungen) analysiert und synthetisiert. Die Ergebnisse wurden im Rahmen eines Syspons-internen Workshops validiert. Ziel der **Auswertung und Validierung der Daten** war es, zusammenfassende Aussagen zu relevanten Untersuchungsaspekten zu treffen.

### **2. Desk Research**

Um weitere Informationen insb. zu GIZ-internen Prozessen zu erhalten, wurde eine **Desk Research elf weiterer Dokumente** durchgeführt. Dazu zählen GIZ-interne Studien, die Checkliste *Evaluability Assessment* und die Prozessbeschreibung und Checkliste für die Prüfung. Hierbei sollte insbesondere beantwortet werden, inwiefern diese Dokumente Schnittstellenfunktionen zwischen Monitoring und Evaluierung einnehmen können, und welche Vorgaben es im Planungsprozess gibt, um Evaluierungen/Evidenzen gezielt bei der Planung neuer Vorhaben abzufragen und einzubringen.

### **3. AV(-V)-Befragung**

Parallel zur *Desk Research* hat die STS Evaluierung die **AVs beziehungsweise AV-Vs aller Vorhaben der Stichprobe mit Folgevorhaben** kontaktiert und **informell per E-Mail gefragt**, inwiefern Empfehlungen der ZPE-Berichte für die Planung der Folgevorhaben genutzt wurden (entsprechend der Dimension Evaluierungs-Planung). Es gab **15 Rückmeldungen**. Die gesammelten Antworten flossen als zusätzliche Datenquelle in die Auswertung ein und wurden als Kriterium zur Auswahl der Deep Dives genutzt.

### **4. Deep Dives**

In einem weiteren Schritt wurden **fünf Deep Dives** mit ausgewählten Vorhaben durchgeführt. Für die vertieften Interviews wurden Vorhaben ausgewählt, die im Rahmen der AV(-V)-Befragung und/oder der qualitativen Auswertung der ZPE-Berichte als potenzielle Beispiele für gute Praktiken insbesondere für die Synchronisierung zwischen Evaluierung und Planung aufgefallen waren. Im Rahmen der fünf Deep Dives wurde ein (Gruppen-)Interview pro Vorhaben geführt, an dem insbesondere der\*die AV beziehungsweise AV-V sowie Monitoringbeauftragte teilnahmen. Ziel der Deep Dives war es unter anderem, vertiefte Kenntnisse zur Qualität der Datenquellen und des WoM, der Abbildung von Querschnittsthemen im Monitoringsystem der Vorhaben, der Rolle von GIZ-internen Dokumenten, der Rolle von Evaluierungen/Evidenzen für Folgevorhaben, sowie der Identifizierung von Erfolgsfaktoren und *Lessons learned* zu gewinnen.

### **5. Interner Syntheseworkshop**

Auf Grundlage der bereits getätigten Auswertungen wurden am Ende der Analysephase die Ergebnisse der QSA in einem **internen Syntheseworkshop** mit allen Expert\*innen des Auswertungsteams reflektiert, vergleichend betrachtet und in der Zusammenschau bewertet. Durch diese Form der Forschertriangulation konnte die Reliabilität bei der Datenauswertung erhöht werden. Darüber hinaus wurden im Syntheseworkshop bereits

erste Handlungsfelder identifiziert und entlang der relevanten Fragestellungen strukturiert. Diese Handlungsfelder wurden mit ersten Handlungsoptionen hinterlegt, die sich aus den Erkenntnissen der Untersuchung ergaben.

### Vorgehen in der Berichtslegungsphase

Ziel der Berichtslegungsphase war es, die Schlussfolgerungen und Handlungsempfehlungen in einem Hauptbericht festzuhalten und die Ergebnisse der QSA PME zu vergemeinschaften.

Im Anschluss an den internen Syntheseworkshop wurden der GIZ die Ergebnisse in Form einer Ergebnispräsentation vorgelegt. Nach der schriftlichen Rückmeldung der GIZ zur Ergebnispräsentation wurde der Entwurf des Hauptberichts vorgelegt. Dieser wurde von der GIZ schriftlich kommentiert. Ziel der Kommentierung war es insbesondere, Rückmeldung zu den Ergebnissen und Schlussfolgerungen der QSA zu erhalten sowie die Handlungsempfehlungen weiterzuentwickeln. Alle Anmerkungen und Anregungen wurden aufgenommen und der GIZ im **finalen Hauptbericht** fristgemäß vorgelegt.

## 3 Ergebnisse der Querschnittsauswertung

In diesem Kapitel werden die zentralen Ergebnisse der QSA PME vorgestellt. Im folgenden Abschnitt wird zunächst auf die Verzahnung zwischen Planung und Monitoring eingegangen. Anschließend wird der Zusammenhang zwischen Monitoring und Evaluierung untersucht. Schließlich wird die Verzahnung zwischen Evaluierung und Planung betrachtet.

### 3.1 Planung - Monitoring

**Ein wesentlicher Anspruch an die Planung der Vorhaben ist, dass diese dazu beiträgt, ein effektives Wirkungsorientiertes Monitoring (WoM) der Vorhaben zu etablieren.** So sollen insbesondere die Wirkungslogik (Wirkungsmodell) und die mit dem Auftraggeber abgestimmten Indikatoren (Wirkungsmatrix) auf der Outcome- und der Output-Ebene das Monitoring der Vorhaben unterstützen. Daher wurde in dieser Querschnittsauswertung die Qualität der evaluierten Projektkonzepte sowie ihr Einfluss auf das Projektmonitoring betrachtet. Die Erkenntnisse zur Synchronisierung zwischen Planung und Monitoring der evaluierten Vorhaben sind dabei nach Wirkungsmodellen, Indikatoren und Datenquellen geordnet.

#### Wirkungsmodelle

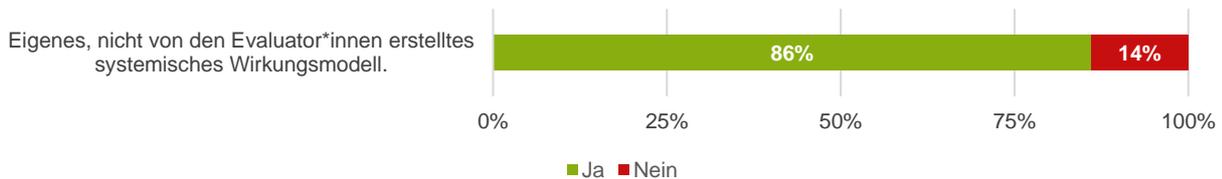
Um ein Wirkungsorientiertes Monitoring in den Vorhaben zu etablieren, ist es entscheidend, ein qualitativ hochwertiges Wirkungsmodell in der Planung zu entwickeln. Hierzu zeigt die qualitative Auswertung, dass die meisten Vorhaben zwar ein eigenes systemisches Wirkungsmodell haben, dieses aber überwiegend vom Evaluatsteam angepasst wird, um dessen Qualität zu erhöhen. Die Mehrheit der Vorhaben haben vor der ZPE bereits ein eigenes, systemisches Wirkungsmodell entwickelt, das Outputs, Outcomes und Impacts aufführt.<sup>3</sup>

---

<sup>3</sup> Aus den ZPE-Berichten ist jedoch nicht ersichtlich, ob diese Wirkungsmodelle in der Planung des Vorhabens oder während der Umsetzung entwickelt wurden. Im zweiten Fall ist zu beachten, dass diese Wirkungsmodelle zu diesem Zeitpunkt bereits durch die im Angebot vereinbarten Wirkungsmatrix etc. eingeschränkt sind.

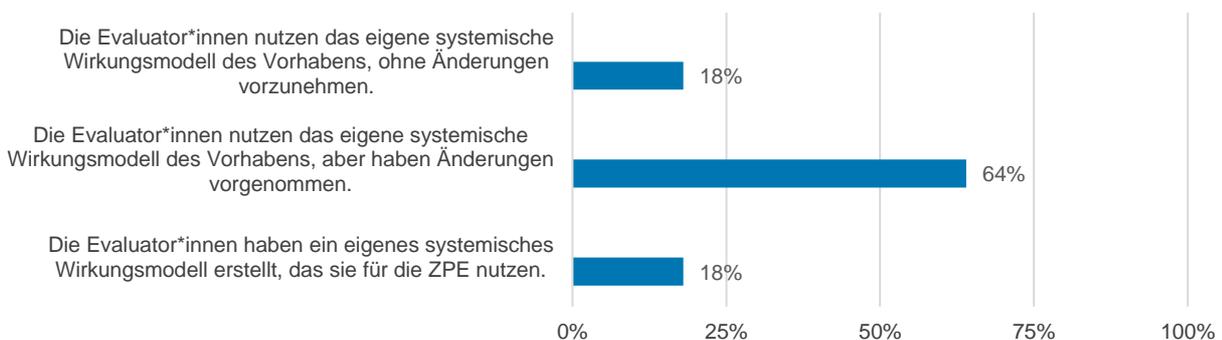
Im Rahmen der ZPE werden diese Wirkungsmodelle von den Evaluationsteams untersucht und ggf. verbessert, angepasst und/oder ergänzt.<sup>4</sup> In der qualitativen Auswertung wurde dabei einerseits deutlich, dass die meisten Vorhaben ein eigenes systemisches Wirkungsmodell haben (siehe Abbildung 2); andererseits war zu erkennen, dass die Evaluationsteams das Wirkungsmodell der Vorhaben zwar überwiegend nutzen, allerdings erst, nachdem sie Änderungen vorgenommen haben (siehe Abbildung 3). Dabei erstellen in zwei von zehn ZPE die Evaluationsteams laut ZPE-Bericht ein komplett neues Wirkungsmodell.

Abbildung 2: Wirkungsmodell der betrachteten Vorhaben



Quelle: Syspons 2022

Abbildung 3: Nutzung der Wirkungsmodelle der betrachteten Vorhaben in den ZPE

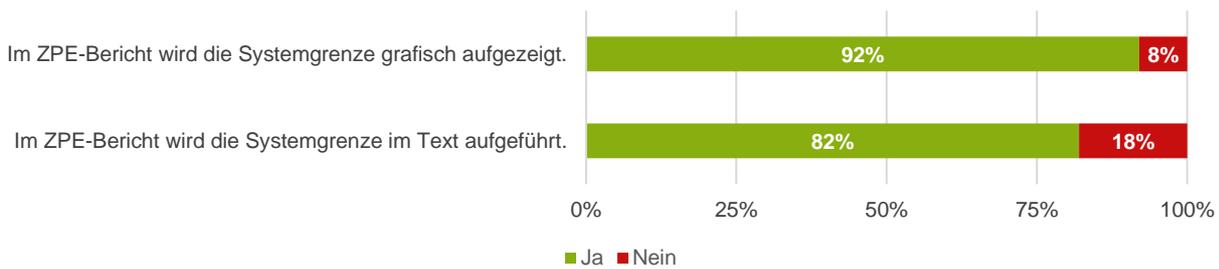


Quelle: Syspons 2022

**Zu einem qualitativ hochwertigen Wirkungsmodell gehört außerdem eine angemessene Systemgrenze, die den Verantwortungsbereich des Vorhabens adäquat aufzeigt. In der qualitativen Auswertung wurde deutlich, dass dies im Großteil der Wirkungsmodelle der Fall ist.** Die Systemgrenze (auch Attributionslücke oder Zuordnungslücke genannt) stellt den direkten Einflussbereich des Vorhabens dar beziehungsweise zeigt auf, ab wann externe Einflüsse so stark werden, dass sie die angestrebten Ziele und/oder Wirkungen der Vorhaben maßgeblich (insbesondere negativ) beeinflussen können. In fast allen ZPE-Berichten ist die Systemgrenze sowohl grafisch im Wirkungsmodell aufgezeigt als auch im Text aufgeführt (siehe Abbildung 4). Grafisch geschieht dies durch die Verortung der Meilensteine, Ziele sowie Wirkungen innerhalb/auf/außerhalb der oft als Kreis dargestellten Systemgrenze. In der qualitativen Auswertung wurde diese Verortungen überprüft, dabei wurde deutlich, dass die Systemgrenze überwiegend adäquat dargestellt ist. Ob die Systemgrenze durch die Evaluationsteams in der ZPE verändert wurde, wird in den ZPE-Berichten allerdings nicht erwähnt.

<sup>4</sup> Diese Revision ist deshalb notwendig, weil die Wirkungsmodelle zentrale Grundlage der Evaluierungen sind. So sind sie der Ausgangspunkt für die Analyse und Bewertung in den Kategorien der OECD/DAC-Kriterien ‚Effektivität‘ und ‚Impact‘: Dafür werden vor der eigentlichen Evaluierung gemeinsam mit dem Projektteam die relevanten Wirkungshypothesen (zum Beispiel Aktivitäten und Meilensteine zu Zielen auf Output- und Outcome-Ebene führen) identifiziert, untersucht und letztlich bestätigt/teilweise bestätigt/nicht bestätigt.

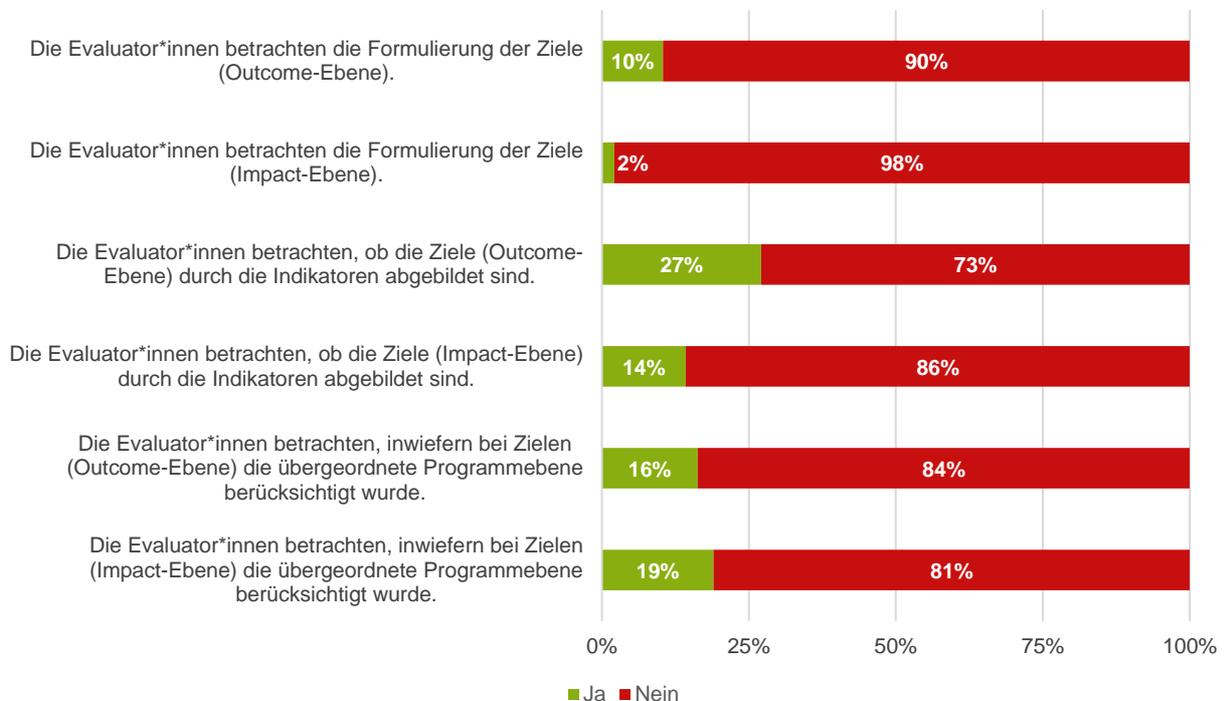
Abbildung 4: Systemgrenze der betrachteten Vorhaben



Quelle: Syspons 2022

**Während der Planungsphase der Vorhaben werden auch dessen angestrebte Ziele und Wirkungen erarbeitet. Dazu zählen einerseits die Ziele auf der Outcome- und der Impact-Ebene selbst und andererseits die Indikatoren, die die Zielerreichung messen sollen. In der QSA wurde daher auch untersucht, inwiefern die Formulierungen der Ziele sowie die Beziehung zwischen Zielen und Indikatoren von den Evaluationsteams der ZPE als gut befunden wurde. Es wurde jedoch deutlich, dass diese von den Evaluationsteams nur selten explizit betrachtet werden.** So wird in nur einem von drei ZPE-Berichten explizit reflektiert, inwiefern die Ziele der Vorhaben in den darunterliegenden Indikatoren abgebildet sind (siehe Abbildung 5). Mit Blick auf die Zielformulierungen (insbesondere auf der Modulziel-, das heißt der Outcome-Ebene) war zu erkennen, dass diese nur von einem von zehn Evaluationsteams im ZPE-Bericht explizit betrachtet wird. Bei der Formulierung der Wirkungen auf Impact-Ebene geschieht dies noch seltener.

Abbildung 5: Untersuchung der Elemente der Wirkungsmodelle der betrachteten Vorhaben in der ZPE



Quelle: Syspons 2022

**Wie eingangs aufgeführt, soll die Planung der Vorhaben dazu beitragen, ein effektives WoM der Vorhaben zu etablieren. Aus der Betrachtung der ZPE-Berichte war der Nutzen der Wirkungsmodelle für das Monitoring der Vorhaben jedoch nicht ersichtlich. Daher wurden die Teilnehmer\*innen der Deep Dives (AVs und Monitoringbeauftragte) gebeten, drei Hypothesen zur Synchronisierung von Planung und Monitoring zu bewerten.**

Zunächst wurde um die Einschätzung einer Hypothese gebeten, die sich nur auf die Planung der Vorhaben konzentriert. Die erste Hypothese lautete: „Gute Wirkungsmodelle sorgen für nachvollziehbare Wirkungshypothesen“ (**Hypothese 1**). Diese Hypothese wurde von allen Befragten **bestätigt**. So wies ein Interviewpartner darauf hin, dass Wirkungsmodelle, die auf einem Prozess beruhen, der das ganze Projektteam miteinbezieht, für ein gemeinsames Verständnis bezüglich der Wirkungserreichung im Projektteam sorgen.

Danach wurden die Teilnehmer\*innen darum gebeten, folgende Hypothese einzuschätzen: „Gute Wirkungsmodelle sorgen für Monitoringdaten, die hilfreich für die Steuerung von Vorhaben sind“ (**Hypothese 2**). Diese Hypothese wurde nur **teilweise bestätigt**: Zwar stimmten drei von fünf Vorhaben zu, zwei Vorhaben stimmten jedoch nicht zu. Letztere wiesen darauf hin, dass Wirkungsmodelle sehr strategisch seien. Sie würden zwar einen guten Horizont vorgeben; jedoch müsse noch viel Arbeit aufgewendet werden, um aus diesen auch Steuerungsinstrumente zu entwickeln.

In der dritten Hypothese wurden die Dimensionen der Planung und des Monitorings zudem mit der Evaluierung verbunden. In diesem Sinn wurden die Teilnehmer\*innen der Deep Dives gebeten, die folgende Hypothese zu bewerten: „Gute Wirkungsmodelle sorgen für Monitoringdaten<sup>5</sup>, die hilfreich für die Evaluierung von Vorhaben sind“ (**Hypothese 3**). Diese Hypothese wurde **überwiegend bestätigt**: So stimmten vier Vorhaben zu und ein Vorhaben dagegen. Aber auch die zustimmenden Projektteams wiesen darauf hin, dass eine Evaluierung tiefergehen würde, und die Monitoringdaten daher nicht die alleinige Datengrundlage sein könnten.

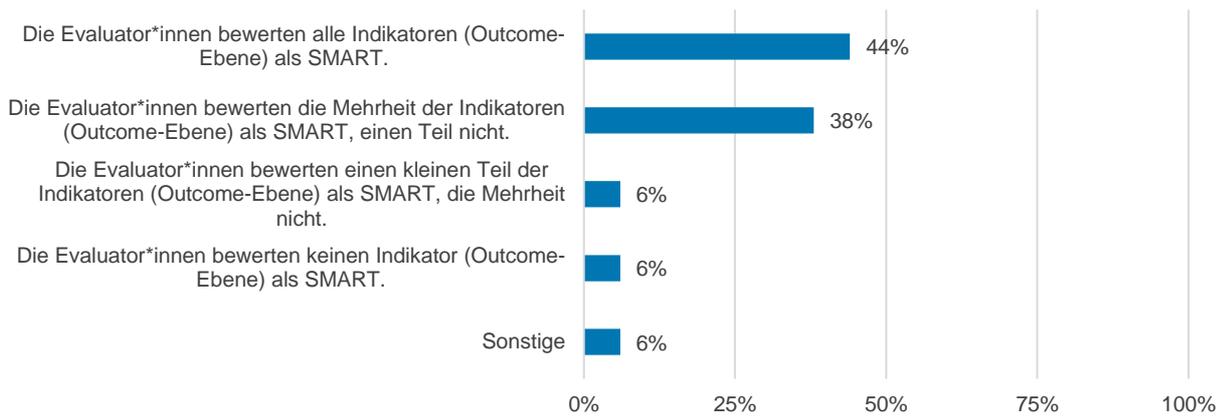
## Indikatoren

**Dass qualitativ hochwertige Indikatoren entscheidend für ein angemessenes WoM sind, ist naheliegend. Die Erstellung von SMARTen (das heißt spezifischen, messbaren, angemessenen, relevanten und zeitgebundenen) Indikatoren in der Planung der Vorhaben ist daher verbindlich vorgegeben. Die qualitative Auswertung zeigt allerdings, dass die Evaluationsteams in der Hälfte der Vorhaben mindestens einen Teil der geprüften Indikatoren kritisieren.** In der Dokumentenanalyse wurde deutlich, dass die Erstellung sowie Prüfung SMARTer Indikatoren verbindlich ist. Dies wird in den Verfahrensinformationen (zum Beispiel *VI083 Qualitätsstandards für Indikatoren*, BMZ 2020c) und Handreichungen (zum Beispiel *Indikatoren: Eine Arbeitshilfe*, GIZ 2014) entsprechend dargestellt und wurde im Rahmen der Gespräche der Deep Dives bestätigt. Die Prüfung der Indikatoren der Outcome-Ebene auf SMART ist auch zentraler Bestandteil der ZPE. Die qualitative Auswertung der ZPE-Berichte zeigt jedoch, dass in sechs von zehn Vorhaben mindestens ein Teil der Indikatoren auf der Outcome-Ebene als nicht-SMART bewertet wird (siehe Abbildung 6). In jedem zehnten Vorhaben betrifft das sogar alle Indikatoren. Es wird insbesondere die fehlende Spezifität der Indikatoren auf der Outcome-Ebene kritisiert, die sich auch negativ auf ihre Messbarkeit auswirkt (siehe Abbildung 7).

---

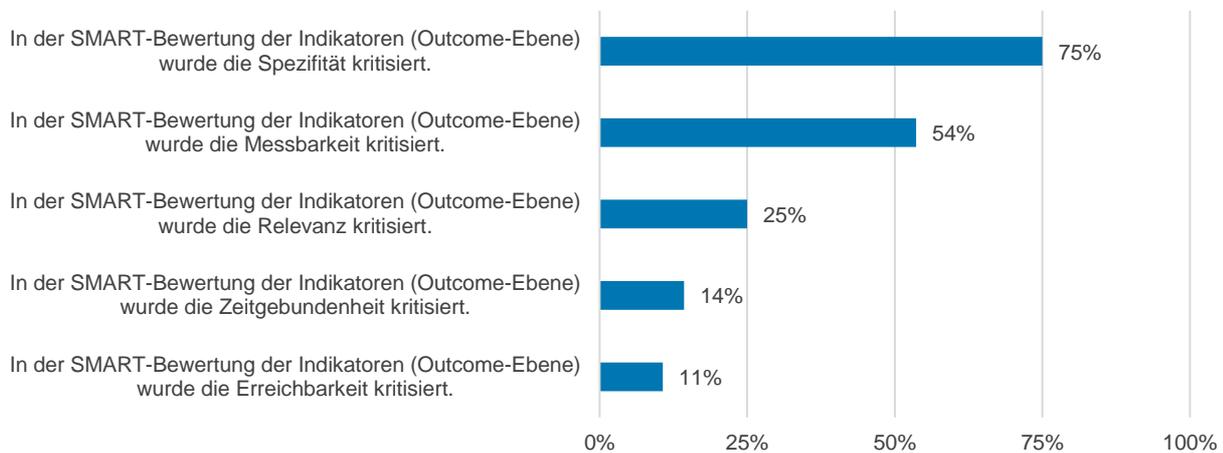
<sup>5</sup> Der Begriff „Monitoringdaten“ ist hier breit gefasst. Es sind alle Daten gemeint, die im Rahmen des Monitorings des Vorhabens erfasst werden. Sie können daher neben der abgestimmten Wirkungsmatrix auch zusätzliche oder weitergehende vom Vorhaben eingeführte Indikatoren oder Informationen umfassen.

Abbildung 6: SMART-Bewertung der Indikatoren der betrachteten Vorhaben in der ZPE



Quelle: Syspons 2022

Abbildung 7: SMART-Anpassung der Indikatoren der betrachteten Vorhaben in der ZPE

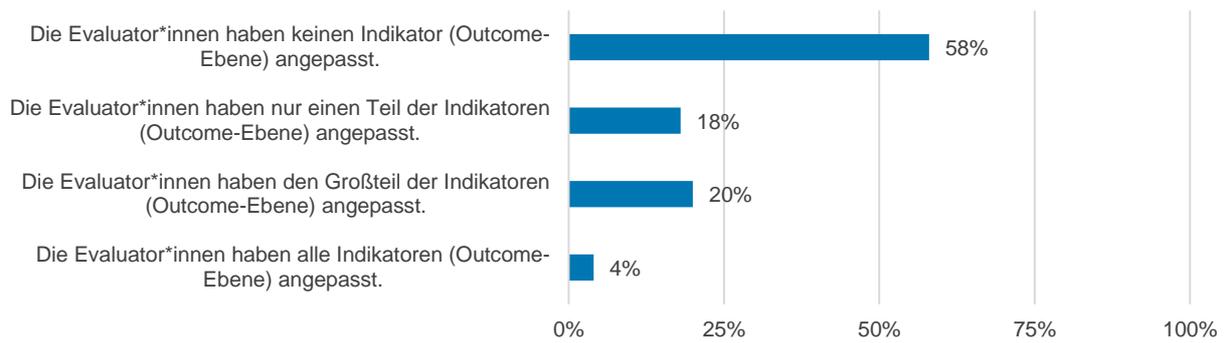


Quelle: Syspons 2022

**Die kritisierten Indikatoren werden von den Evaluationsteams angepasst, um deren Qualität zu erhöhen. In diesem Sinne wurden in fast der Hälfte der ZPE-Berichte Indikatoren angepasst, damit diese monitorbar sowie evaluierbar werden.** In vier von zehn Vorhaben nehmen Evaluatord\*innen in der ZPE die Änderung von mindestens einem Teil der Indikatoren vor<sup>6</sup> (siehe Abbildung 8). In fast jedem vierten der betrachteten Vorhaben wurde dabei der Großteil der Indikatoren und teils sogar alle Indikatoren verändert, damit diese SMART werden. Die qualitative Betrachtung hat gezeigt, dass die Evaluationsteams laut ZPE-Berichten meist die Spezifität anpassen, ebenso wie die Messbarkeit der Indikatoren (siehe Abbildung 9).

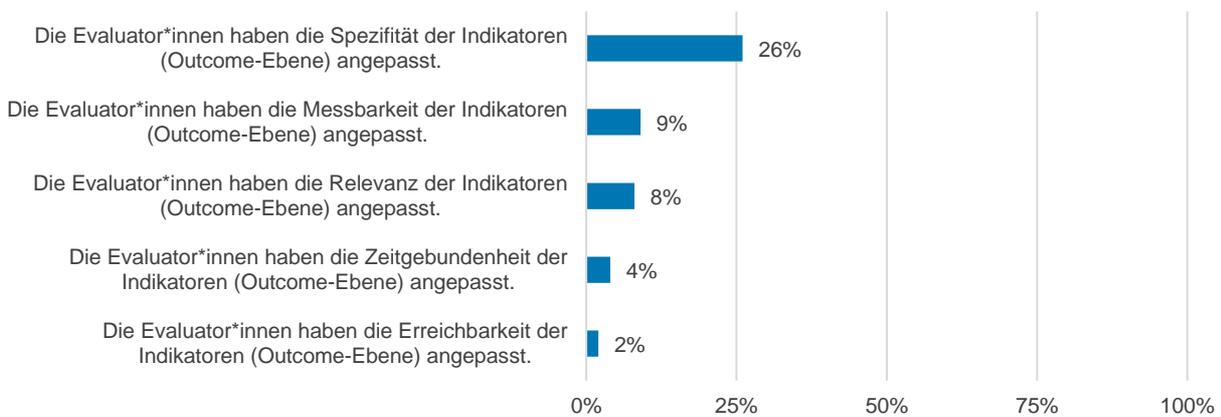
<sup>6</sup> Diese Revision ist deshalb notwendig, weil die Indikatoren zentrale Grundlage der Evaluierungen sind. So sind sie der Ausgangspunkt für die Analyse und Bewertung in den Kategorien der OECD/DAC-Kriterien ‚Effektivität‘ und ‚Effizienz‘. Damit die Indikatoren von den Evaluierungsteams für diese Bewertung genutzt werden können, prüfen die Evaluierungsteams vor der eigentlichen Evaluierung, ob die Indikatoren spezifisch, messbar, angemessen, relevant und zeitgebunden (d.h. SMART) sind, und verändern sie gegebenenfalls.

Abbildung 8: Anzahl der SMART-angepassten Indikatoren der betrachteten Vorhaben in der ZPE



Quelle: Syspons 2022

Abbildung 9: SMART-Kategorie, die in der ZPE der betrachteten Vorhaben angepasst wurde<sup>7</sup>



Quelle: Syspons 2022

## Datenquellen

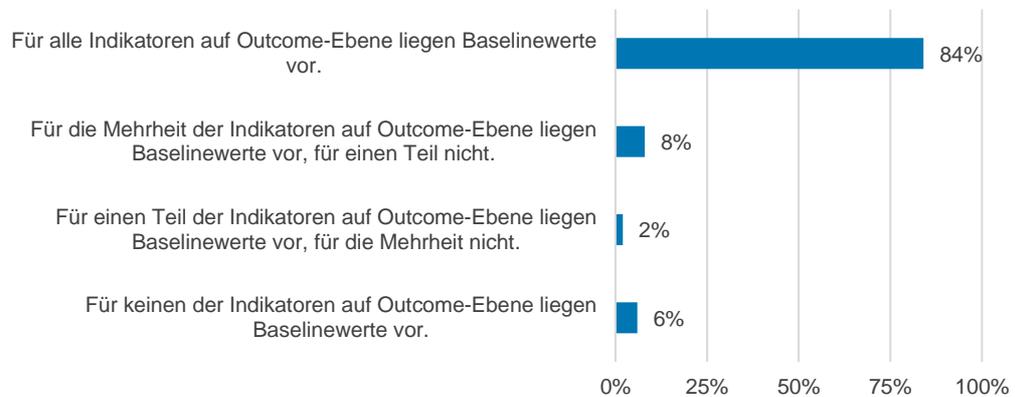
**Damit das WoM der Vorhaben qualitativ hochwertig ist, müssen adäquate Datenquellen genutzt werden. Damit ist unter anderem gemeint, dass Vorhaben verschiedene Datenquellen (zum Beispiel Daten der Partnerorganisationen, nationale Statistiken) in ihrem Monitoringsystem kombinieren.** In der qualitativen Auswertung lag der Fokus dabei zunächst auf den Baseline-, das heißt den Ausgangswerten der Indikatoren, doch waren auch Zielwerte Teil der Untersuchung. Das laufende Monitoring wurde dabei im Rahmen der Synchronisierung zwischen Monitoring und Evaluierung analysiert (siehe Kapitel 3.2).

**Während der Planung werden üblicherweise die Baselinewerte der Indikatoren festgelegt, nur in Ausnahmefälle werden diese im Projektverlauf nachgereicht. Laut der qualitativen Auswertung ist der Großteil der Baselinewerte dabei zunächst nachvollziehbar gewählt. Hinweise auf die für die Erhebung der Baselinewerte genutzten Datenquellen gibt es dabei aber nur in der Hälfte der Vorhaben, meist sind es Primär- und/oder Daten der Partnerorganisationen.** Im Großteil der Vorhaben liegen Baselinewerte für die Indikatoren auf der Outcome-Ebene vor (siehe Abbildung 10). Auf Grundlage der Untersuchung und Einschätzung der Auswerter\*innen waren in acht von zehn Vorhaben mindestens ein Teil der Baselinewerte nach-

<sup>7</sup> Während Abb. 6 und 7 die Bewertung der Indikatoren nach SMART-Kriterien thematisieren, bilden Abb. 8 und 9 ab, inwiefern die Indikatoren letztlich tatsächlich angepasst wurden.

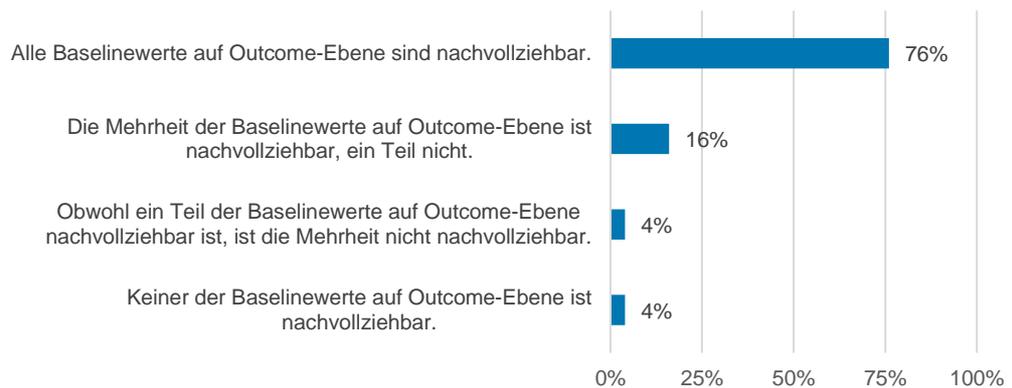
vollziehbar, in sechs von zehn Vorhaben war dies für alle Baselinewerte der Fall (siehe Abbildung 11). Hinsichtlich der Erfassung der Baselinewerte (Datenquellen) wurde deutlich, dass nur in der Hälfte der Vorhaben eine Baselinestudie durchgeführt und/oder Daten der nationalen Partnerorganisationen/Stakeholder und/oder Sekundärdaten genutzt wurden (54 %). Die genaue Verteilung ist der Abbildung 12 zu entnehmen.

Abbildung 10: Anzahl der vorliegenden Baselinewerte für die Indikatoren der untersuchten Vorhaben



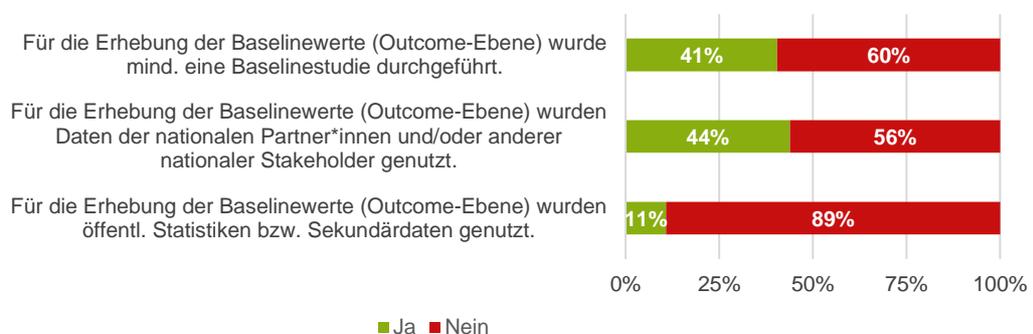
Quelle: Syspons 2022

Abbildung 11: Nachvollziehbarkeit der Baselinewerte für die Indikatoren der untersuchten Vorhaben



Quelle: Syspons 2022

Abbildung 12: Art der Datenerhebung der Baselinewerte für die Indikatoren der untersuchten Vorhaben

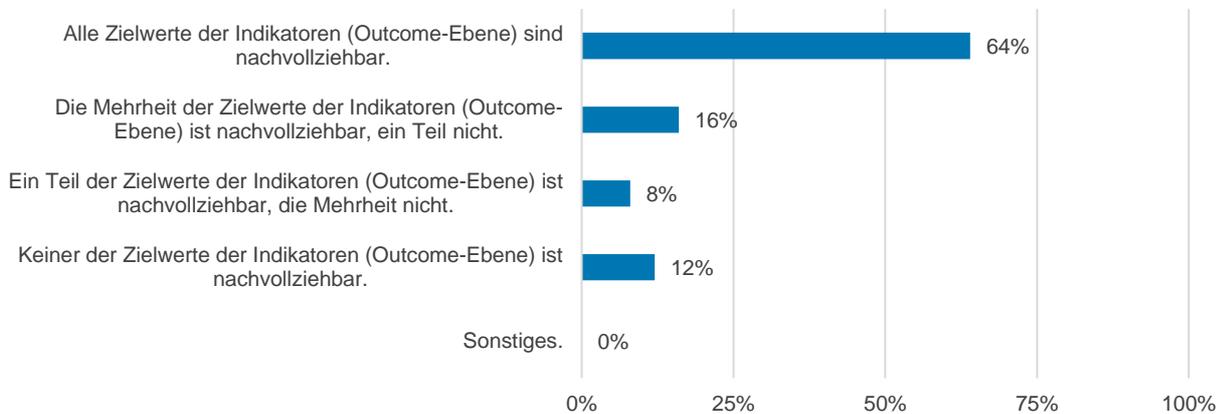


Quelle: Syspons 2022

**Für ein angemessenes WoM müssen auch die Zielwerte der Indikatoren nachvollziehbar gewählt sein. In der qualitativen Auswertung wurde deutlich, dass für alle Indikatoren auf der Outcome-Ebene Zielwerte vorliegen und dass diese in der Mehrheit der betrachteten Vorhaben nachvollziehbar sind. Bei**

der Untersuchung der ZPE-Berichte durch die Auswerter\*innen war in fast neun von zehn Berichten mindestens ein Teil der Zielwerte der Indikatoren auf der Outcome-Ebene nachvollziehbar (siehe Abbildung 13). In mehr als sechs von zehn Vorhaben war dies für alle Zielwerte der Indikatoren auf der Outcome-Ebene der Fall. Welche Datenquellen für die Zielwerte genutzt werden, ist in der Betrachtung des laufenden Monitorings aufgeführt (siehe Kapitel 3.2 beziehungsweise Abbildung 14).

Abbildung 13: Nachvollziehbarkeit der Zielwerte der Indikatoren der untersuchten Vorhaben



Quelle: Syspons 2022

### Erfolgs- und Misserfolgskriterien

**Zusammenfassend ist in der qualitativen Auswertung zu erkennen, dass sich die bisherigen Vorgaben/Handreichungen sowie das Wirkungsmodell als Erfolgsfaktoren für eine Synchronisierung zwischen Planung und Monitoring herausgestellt haben.** So wurde in den Interviews der Deep Dives oft auf die Vorgaben/Handreichungen zur Erstellung der Indikatoren hingewiesen: Die Vorgaben/Handreichungen hätten einen positiven Einfluss darauf gehabt, dass SMARTe Indikatoren erstellt worden wären. In den Vorgaben/Handreichungen wird dabei einerseits festgehalten, dass (1) SMARTe Indikatoren erstellt werden und (2) dies in der Prüfung des Vorhabens auch geprüft werden müsse(n); andererseits werden in den Handreichungen auch Informationen/Anleitungen zur Formulierung SMARTer Indikatoren zur Verfügung gestellt. Es wurde deutlich, dass das Wirkungsmodell ein weiterer Erfolgsfaktor ist: Laut Monitoringbeauftragten mache dies deutlich, welche Daten erhoben werden sollten, und unterstütze ein geteiltes Verständnis für Ziele und Zielerreichung in den Projektteams. Allerdings muss ebenfalls darauf hingewiesen werden, dass die erstellten Indikatoren in der SMART-Bewertung in den ZPE eher kritisch bewertet wurden.

**Im Gegensatz dazu haben sich in der Auswertung die für das Monitoring bereitgestellten, zu geringen, Ressourcen als Misserfolgskriterien für eine Synchronisierung zwischen Planung und Monitoring herausgestellt.** So wurde in den Interviews sowie in ausgewählten ZPE-Berichten darauf hingewiesen, dass im Vorhaben angemessene Ressourcen für ein erfolgreiches Monitoring während der Umsetzung des Vorhabens eingeplant werden müssten. Ein erfolgreiches Monitoring müsse entsprechend vorbereitet, aufgesetzt, begleitet, durchgeführt und auch ausgewertet werden.

## 3.2 Monitoring - Evaluierung

**Das Monitoring der Vorhaben soll nicht nur zur Steuerung des Vorhabens, sondern auch zu dessen Evaluierung beitragen.** Daher wurde in dieser Querschnittsauswertung betrachtet, inwiefern die Monitoringssysteme der Vorhaben eine geeignete Grundlage für ihre spätere Evaluierung, insbesondere ZPE, darstellen. Die Erkenntnisse zur Synchronisierung zwischen Monitoring und Evaluierung der ausgewählten Vorhaben sind dabei entsprechend dem Monitoring der Vorhaben, dem Monitoring der Wirkungshypothesen, den Erkenntnissen zu Beiträgen der Vorhaben auf der Impact-/Programm-Ebene und der Bewertung des WoM der Vorhaben sortiert. Zudem wurde auch das von der Stabsstelle Evaluierung entwickelte *Evaluability Assessment* (GIZ n.d.) betrachtet.

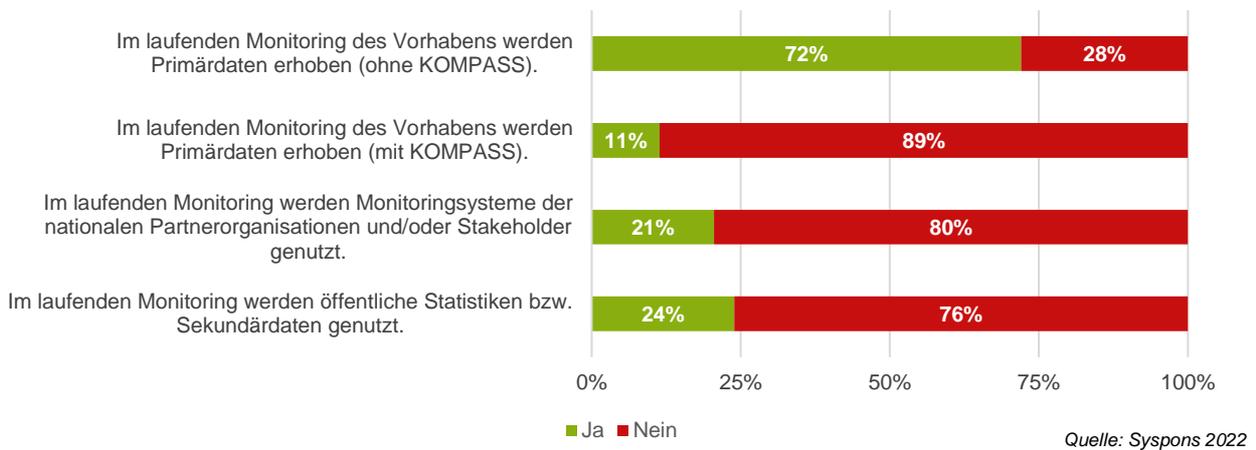
### Monitoring der Vorhaben

**Damit das Monitoring des Vorhabens dessen spätere Evaluierung unterstützen kann, ist es für das Evaluationsteam wichtig, auf ein qualitativ hochwertiges WoM des Vorhabens zurückgreifen zu können. Um die Erkenntnisse der qualitativen Auswertung der ZPE-Berichte einzuordnen, wurden zunächst mithilfe einer Dokumentenanalyse Vorgaben des BMZ sowie der GIZ identifiziert. So wurde deutlich, dass in den Verfahrensinformationen und Handreichungen explizite Hinweise zum WoM aufgeführt sind.** Diese Hinweise thematisieren dabei sowohl den Aufbau des Monitoringsystems als auch weitere Aspekte, wie zum Beispiel die Regelmäßigkeit der Datenerhebungen im laufenden Monitoring. Von Seiten des BMZ fokussieren die Anweisungen auf die abzubildenden Indikatoren: So wird in der Verfahrensinformation VI083 darauf hingewiesen, dass im Monitoring verschiedene Ebenen der Zielerreichung, die Sensibilität/Reaktionszeit der Indikatoren sowie steuerungsrelevante Indikatoren berücksichtigt werden sollen (BMZ 2020c). Die Arbeitshilfe der GIZ zu Indikatoren führt zudem praktische Aspekte des Monitorings auf (GIZ 2014). So wird hier zum Beispiel auf die Überprüfung der Indikatoren in Meilenstein-Intervallen verwiesen und thematisiert, unter welchen Voraussetzungen Sekundärdaten aus Partnersystemen und/oder nationalen statistischen Systemen genutzt werden sollten. In den untersuchten Dokumenten gab es zudem konkrete Verweise auf weitere Unterlagen, wie die „Anleitung zum Aufbau und zur Nutzung eines WoM-Systems der GIZ“.

Darauf aufbauend wurden Angaben zum Monitoring der Vorhaben (und wie diese im Rahmen der ZPE genutzt wurden) in den ZPE-Berichten der Evaluationsteams untersucht. Hier muss jedoch nochmal deutlich darauf verwiesen werden, dass **nur explizite Angaben** der Evaluationsteams erfasst wurden.

**So wurde in der Dokumentenanalyse zwar die Empfehlung für das WoM der Vorhaben identifiziert, verschiedene Datenquellentypen zu nutzen. In der qualitativen Betrachtung der ausgewählten ZPE-Berichte wurde aber deutlich, dass im Monitoringsystem der Vorhaben Primärdaten die zentrale Datenquelle sind. Insbesondere die offene Perspektivenerfassung für Partnerorganisationen und Zielgruppen wird selten genutzt.** In der Mehrheit der Vorhaben (83%) werden somit Primärdaten erhoben, das heißt relevante Daten werden vom Projektteam selbst und/oder den Partnerorganisationen der Vorhaben eingepflegt. Dies geschieht jedoch nur selten unter Verwendung der offenen Perspektivenerfassung ‚KOMPASS‘, mit der Partnerorganisationen und Zielgruppen eingebunden und die Wirkungsmessung anhand der Indikatoren ergänzt werden soll(en). Weiterhin wurden nur in 36% der Vorhaben entweder Daten der nationalen Partner/Stakeholder und/oder öffentliche Statistiken/Sekundärdaten genutzt. Die genaue Verteilung ist Abbildung 14 zu entnehmen. Im Rahmen der Datenerhebung, insbesondere der Deep Dives, wurden zudem Beispiele für gute Praktiken identifiziert, die in Abbildung 15 aufgeführt sind.

Abbildung 14: Art der Datenerhebung im laufenden Monitoring der untersuchten Vorhaben



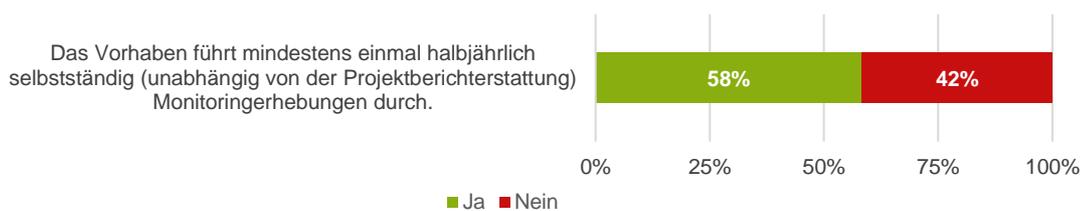
Kasten 1: Beispiele für gute Praktiken: Art der Datenerhebung im laufenden Monitoring

## Beispiele für gute Praktiken

- **Berufliche Bildung und Jugendbeschäftigung III** (PN 2016.2100.2) (Deep Dive): Das Vorhaben bezieht neben der eigenen Primärdatenerhebung für die Projektfortschrittsberichterstattung weitere Sekundärdaten ein, zum Beispiel Statistiken der nationalen Arbeitsagentur, Berichte der Weltbank.
- **Cities Finance Facility** (PN 2018.2102.4) (Deep Dive): Die zentrale Datenquelle sind eigene Primärdaten (in Auftrag gegebene Studien, eigene Maßnahmen). Weiterhin werden Daten direkt von den Partnerorganisationen des Vorhabens erhoben. Das Vorhaben bezieht sich in seinem Monitoring zudem auch auf die gesetzlichen Grundlagen.

**Außerdem wird Vorhaben empfohlen, die Werte der Indikatoren in regelmäßigen Intervallen zu aktualisieren, um diese zur Steuerung nutzen zu können. Regelmäßige Monitoringerhebungen vereinfachen zudem die Nachverfolgung der Umsetzung des Vorhabens und damit dessen Evaluierung. Laut den ZPE-Berichten führt dennoch (nur) etwas mehr als die Hälfte der Vorhaben regelmäßig Monitoringerhebungen durch.** Dies in sechs von zehn ZPE-Berichten der Fall (siehe Abbildung 16). Dabei wurden von den Auswerter\*innen nur jene Monitoringerhebungen erfasst, die vom Vorhaben selbstständig (das heißt unabhängig von ihrer Projektberichterstattung) mindestens einmal halbjährlich durchgeführt wurden. Beispiele für gute Praktiken, die im Rahmen der qualitativen Auswertung und der Deep Dives erfasst wurden, sind in Abbildung 17 aufgeführt.

Abbildung 15: Regelmäßigkeit der selbstständig durchgeführten Monitoringerhebungen der untersuchten Vorhaben

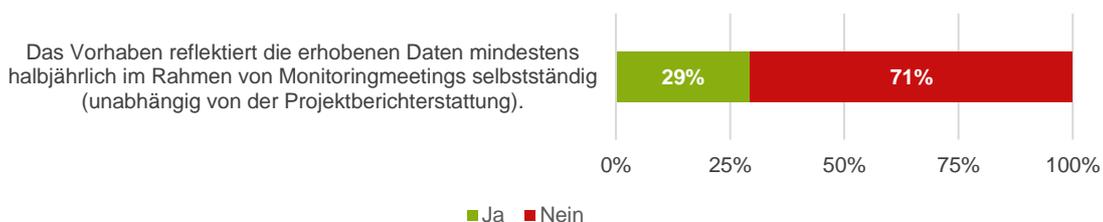


## Beispiele für gute Praktiken

- **Programm zur Förderung lokaler Regierungsführung (PN 2015.2159.0) (Deep Dive):** Das Vorhaben hat vor Projektbeginn einen Aktivitätenplan definiert, in dem Zeiträume und Ressourcen festgehalten wurden. Relevante Daten wurden durch verschiedene Monitoringbeauftragte (es wurde ein\*e Beauftragte\*r pro sog. Prozesscluster ernannt) erhoben und ins Online-Monitoring-Tool eingespeist.
- **Anpassung landwirtschaftlicher Anbaumethoden an den Klimawandel und Stabilisierung der Lebensgrundlagen, Südsudan (PN 2012.9830.6):** Unter der Leitung des\*r Monitoringbeauftragten, führte das Vorhaben regelmäßige Monitoringerhebungen der Aktivitäten vor Ort durch. Das Projektteam wurde dafür in der Nutzung von simplen Vorlagen geschult, die unter anderem Arbeitspläne und Ergebnisse enthielten. Diese Vorlagen wurden wöchentlich ausgefüllt und am Monatsende berichtet.

Das WoM der Vorhaben kann dabei auch von den Projektteams selbst genutzt werden, um eigenständig in Monitoringmeetings über den Verlauf des Vorhabens zu reflektieren. Die qualitative Auswertung der ZPE-Berichte zeigt allerdings, dass in nur jedem dritten Bericht von regelmäßigen Monitoringmeetings der Vorhaben berichtet wird. So war dies in drei von zehn Berichten der Fall (siehe Abbildung 18). Dabei wurden nur jene Fälle erfasst, in denen das Vorhaben selbstständig (das heißt unabhängig von ihrer Projektberichterstattung) mindestens einmal halbjährlich Monitoringdaten im Rahmen von Meetings reflektiert. Auch hier wurden bei der qualitativen Auswertung und den Deep Dives Beispiele für gute Praktiken erfasst. Diese sind in Abbildung 18 aufgeführt.

Abbildung 16: Regelmäßigkeit der selbstständig durchgeführten Monitoringmeetings in den untersuchten Vorhaben



Quelle: Syspons 2022

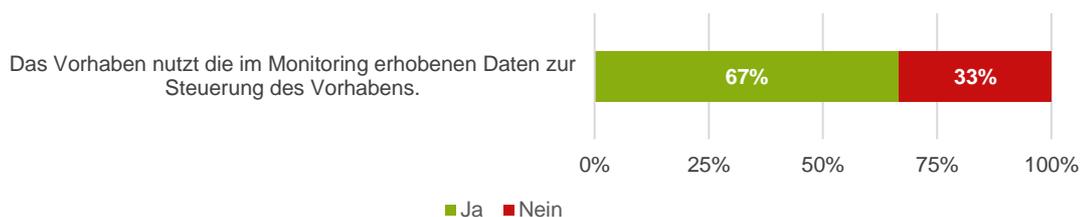
## Beispiele für gute Praktiken

- **Programm zur Förderung lokaler Regierungsführung (PN 2015.2159.0) (Deep Dive):** Das Vorhaben hat vor Projektbeginn einen Aktivitätenplan definiert (siehe Kasten 2). Alle 45 Tage werden Daten, die durch die Monitoringbeauftragten erhoben werden, sowie mögliche Hindernisse und notwendige Revisionen diskutiert.
- **Anpassung landwirtschaftlicher Anbaumethoden an den Klimawandel und Stabilisierung der Lebensgrundlagen, Südsudan (PN 2012.9830.6):** Der auf der Monitoringvorlage basierende, wöchentlich erfasste Projektfortschritt wurde in wöchentlichen Treffen mit Projektpartnern berichtet und reflektiert.

Wie eingangs erwähnt, wird in der BMZ-Verfahrensinformation darauf hingewiesen, dass die vom BMZ finanzierten Vorhaben der GIZ auch steuerungsrelevante Indikatoren wählen sollen. So sollen diese aufzeigen, ob Entscheidungen in Bezug auf Anpassungen notwendig sind (BMZ 2020c). **Im Rahmen der qualitativen Auswertung der ausgewählten ZPE-Berichte wird dabei deutlich, dass der Großteil der Vorhaben die erhobenen Daten zur Steuerung nutzt.** So wurde in der Auswertung der ZPE-Berichte deutlich, dass dies in sieben von zehn Vorhaben geschieht (siehe Abbildung 20). Beispiele für gute Praktiken sind dabei in Abbildung 21 aufgeführt.

Um die Rolle des Monitoringsystems für das Vorhaben einzuschätzen, wurden die Teilnehmer\*innen der Deep Dives gebeten, eine Hypothese zur Verbindung zwischen der Qualität des Monitoringsystems sowie der Nutzung des Monitorings in der Steuerung der Vorhaben zu bewerten. Die Hypothese lautete dabei: **„Wenn das Monitoringsystem gut ist, werden die Monitoringdaten in der Steuerung von Vorhaben genutzt.“ (Hypothese 4)**. Diese Hypothese wurde von allen fünf befragten Projektteams **bestätigt**. Sie wiesen im Rahmen ihrer Zustimmung aber auch darauf hin, dass ein gutes WoM kein Garant für die Nutzung sei, aber eine gute Basis. So seien die Projektteams, wenn sie ein gutes WoM haben, auch motiviert, dies zu nutzen.

Abbildung 17: Nutzung der Monitoringdaten zur Steuerung der untersuchten Vorhaben



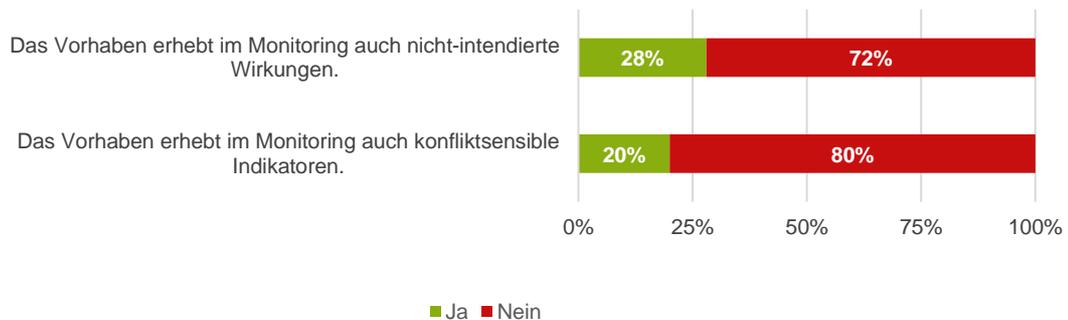
Quelle: Syspons 2022

## Beispiele für gute Praktiken

- **Verbesserung der öffentlichen Finanzen** (PN 2015.2087.3) (*Deep Dive*): Das Monitoringsystem wurde vom Vorhaben auch zur Steuerung genutzt. So wurde der Projektfortschritt betrachtet und entschieden, in welchen Handlungsfeldern keine Aktivitäten mehr durchgeführt werden sollten, um sich auf andere Felder mit geringerem Fortschritt zu fokussieren.
- **Cities Finance Facility** (PN 2018.2102.4) (*Deep Dive*): Das Vorhaben nutzte das eigene Monitoring nach eigenen Angaben zur Steuerung, insbesondere durch den Abgleich der Daten mit den Operationsplänen mit den Partnerstädten. Das Monitoringtool beinhaltete zudem ein Ampelsystem zur Risikobewertung, in dem das Projektteam auf für die Steuerung relevante Informationen wie das Ausscheiden von Mitarbeiter\*innen oder Veränderungen in den Partnerstädten hinweisen konnte.

**In der Evaluierung der Vorhaben wird ebenfalls geprüft, inwiefern diese zu nicht-intendierten Wirkungen beigetragen haben. Auch in dieser Untersuchung soll das Evaluationsteam möglichst vom Monitoringsystem des Vorhabens unterstützt werden. In der qualitativen Auswertung wurde dabei deutlich, dass weniger als die Hälfte der Vorhaben nicht-intendierte Wirkungen und/oder konflikt sensible Daten erfasst.** So war dies in etwa jedem dritten Vorhaben der Fall. Dabei wurde deutlich, dass dies jedoch meist nicht in Form von Indikatoren, sondern in Form von Informationen, die zusätzlich erhoben und reflektiert werden, geschieht. Diese Einschätzung wurde im Rahmen der Deep Dives von den Vorhaben bestätigt.

Abbildung 18: Erhebung von nicht-intendierten Wirkungen im Monitoring der untersuchten Vorhaben



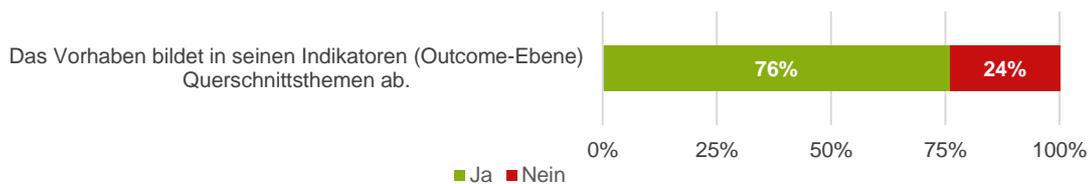
Quelle: Syspons 2022

## Beispiele

- **Nachhaltiges Ressourcenmanagement im Fokus des Klimawandels** (PN 2014.2168.4): Ein Indikator erfasst die Integration von Organisationen der indigenen Bevölkerungsgruppen (positive nicht-intendierte Wirkung des Vorhabens).
- **Cities Finance Facility** (PN 2018.2102.4) (Deep Dive): Das Erheben von nicht-intendierten Wirkungen und/oder konfliktensiblen Indikatoren hat eine Rolle gespielt, da das Vorhaben in Ländern beziehungsweise Städten implementiert wird, in denen es Konflikte gibt. Das Vorhaben misst deshalb zum Beispiel die Resilienz beziehungsweise die Anpassung der Gesellschaft. Das wird nach eigenen Angaben nicht durch Indikatoren abgedeckt, sondern durch ein Risikomanagementsystem. Indirekt, durch die Berichte der Partnerstädte, wird zudem Information zur Umsiedlung der Stadtbevölkerung erhoben.

Um die Fragestellungen der ZPE und Erarbeitung von Empfehlungen zu unterstützen, brauchen die Evaluationsteams zusätzlich Zugang zu Informationen über Querschnittsthemen, wie zum Beispiel Gender, Digitalisierung oder die Agenda 2030. Im besten Fall kann das Evaluationsteam dafür auch auf Daten des Monitoringsystems zurückgreifen. Die QSA hat gezeigt, dass die meisten Monitoringsysteme der Vorhaben Querschnittsthemen, insbesondere Gender, abbilden. So bildeten drei von vier der ausgewählten Vorhaben Querschnittsthemen in ihren Indikatoren auf der Outcome-Ebene ab (siehe Abbildung 24). In den meisten Fällen handelt es sich dabei um gender-disaggregierte Zielwerte, gefolgt von der Agenda 2030 / den SDGs. Andere betrachtete Querschnittsthemen waren unter anderem die Zusammenarbeit mit der Wirtschaft und Digitalisierung. Beispiele sind in Abbildung 25 dargestellt.

Abbildung 19: Abbildung von Querschnittsthemen in den Indikatoren der untersuchten Vorhaben



Quelle: Syspons 2022

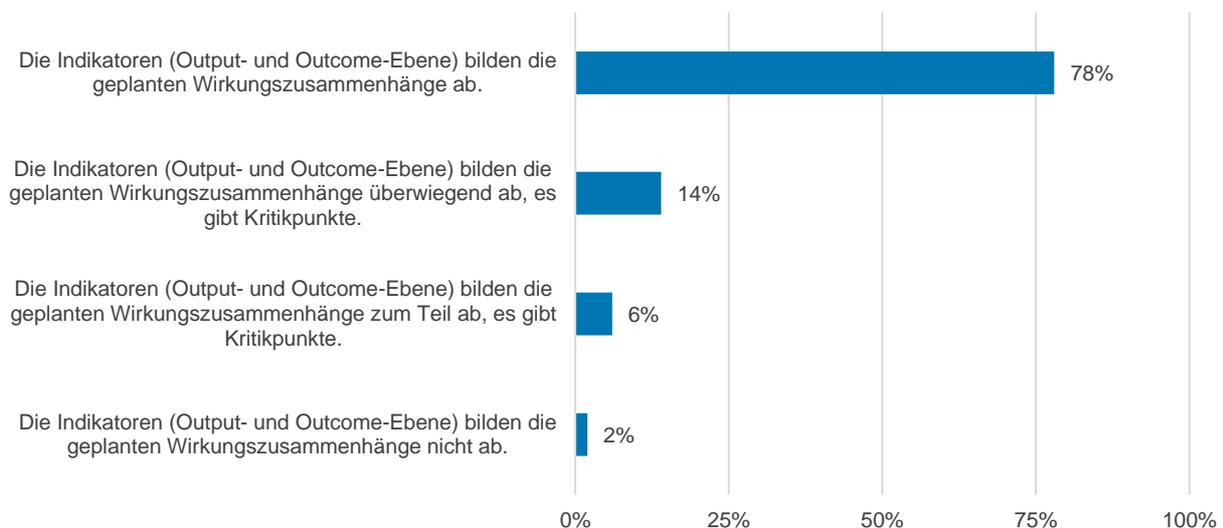
## Beispiele

- **Pilotvorhaben Lernen für die Rückkehr** (PN 2016.2180.4): Modulzielindikator 2: „100 von 140 Nutzer\*innen (davon 30% Frauen) der Zusatzangebote bestätigen, dass die Qualifizierungsmaßnahmen ihren Bedürfnissen entsprechen.“
- **Digitale Lösungen für nachhaltige Entwicklung** (PN 2017.2024.2): Modulzielindikator 2: „Für acht SDG-relevante Lösungen, die im Digitalisierungszentrum entwickelt wurden und vor allem Frauen und Jugendlichen zugutekommen, wurden in den zuständigen Fachministerien Umsetzungs- und Finanzierungspläne erstellt“.

### Monitoring der Wirkungshypothesen

Eine zentrale Anforderung an die ZPE ist, dass Evaluationsteams ausgewählte Wirkungshypothesen überprüfen. Auch hier sollen sie auf Daten des Monitoringsystems des Vorhabens zurückgreifen können. Die qualitative Auswertung hat hier gezeigt, dass der Großteil der Indikatoren auf der Output- und der Outcome-Ebene die geplanten Wirkungszusammenhänge indirekt abbilden. Dadurch findet ein indirektes Monitoring der Wirkungshypothesen statt. Basierend auf den ZPE-Berichten werden die Wirkungshypothesen der Vorhaben insbesondere durch die Erfassung der Indikatoren<sup>8</sup> und eher selten über andere Instrumente wie die offene Perspektivenerfassung (KOMPASS) erhoben. Die Indikatoren von neun von zehn Vorhaben bilden dabei mindestens teilweise die geplanten Wirkungszusammenhänge auf der Output- und der Outcome-Ebene ab (siehe Abbildung 26 und 27).

Abbildung 20: Abbildung der Wirkungshypothesen in den darunterliegenden Indikatoren der untersuchten Vorhaben



Quelle: Syspons 2022

<sup>8</sup> Im Monitoringsystem werden die zu erreichenden Ziele auf der Output-, der Outcome- und gegebenenfalls der Impact-Ebene erfasst. Wenn die dazugehörigen Indikatoren die Ziele adäquat abbilden, findet dadurch ein indirektes Monitoring der Wirkungshypothese (also "wenn Output X, dann Outcome X") statt. Um die Zusammenhänge qualitativ zu bestätigen, müssen die Evaluationsteams die Information der Monitoringsysteme durch eine Dokumentenanalyse, Interviews und Fokusgruppengespräche ergänzen.

## Beispiel

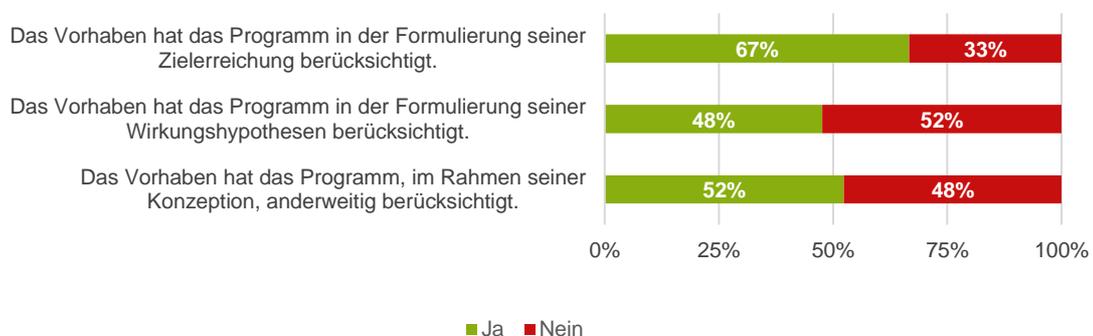
- **Cities Finance Facility (2018.2102.4):**
  - Modulzielindikator 4: „11 Lernformate (zum Beispiel Workshops, Webinare, Onlinetutorials, Konferenzen und Austauschplattformen), die auf CFF-Erfahrungen basieren, werden von 70% der Teilnehmer\*innen aus Städten (50% aus Nicht-C40-Städten), Politiker\*innen, und Entwicklungsorganisationen als nützlich befunden, um für Klimaschutzprojekte in Städten der ganzen Welt Finanzierung zu erhalten.“
  - Wirkungshypothese 2: „Der Austausch von Wissen, das bei der Planung von Investitionen in Klimaschutzprojekte erworben wurde, und der Zugang zu Finanzierung durch Peer-to-Peer-Lernen und Netzwerke und Partnerschaften von CFF-Akteuren über CFF-Partnerstädte hinaus wird dazu beitragen, das Fachwissen zu verbessern, das politische Bewusstsein zu schärfen und Ressourcen zu mobilisieren, damit mehr Städte nachhaltige Investitionen in städtische Klimaschutzmaßnahmen tätigen können.“

### Beiträge auf der Impact-/Programm-Ebene

Die betrachteten Vorhaben können entweder alleine stehen oder zu einem übergeordneten Programm gehören. Ist letzteres der Fall, so soll in der ZPE ebenfalls geprüft werden, inwiefern die Vorhaben zur Erreichung der Ziele auf der Programmebene (das heißt der Impactebene) beitragen. Auch hierzu soll das Monitoringsystem der Vorhaben Daten liefern.

Die qualitative Auswertung zeigt, dass der Großteil der Vorhaben, die einem Programm angehören, dieses in der Formulierung seiner Ziele und/oder Wirkungshypothesen berücksichtigt haben. Nur vereinzelt gibt es jedoch in den ZPE-Berichten Hinweise auf die Indikatoren auf der Impact-Ebene oder eine andere Einbindung in das WoM-System. So haben fast sieben von zehn Vorhaben, die einem Programm angehören<sup>9</sup>, dieses Programm in der Formulierung seiner Zielerreichung berücksichtigt (siehe Abbildung 28). Etwas weniger, nämlich ungefähr fünf von zehn Vorhaben, die einem Programm angehören, haben dieses in der Formulierung seiner Wirkungshypothesen und/oder anderweitig in seiner Konzeption berücksichtigt.

Abbildung 21: Berücksichtigung der Programm-Ebene in den betrachteten Vorhaben



Quelle: Syspons 2022

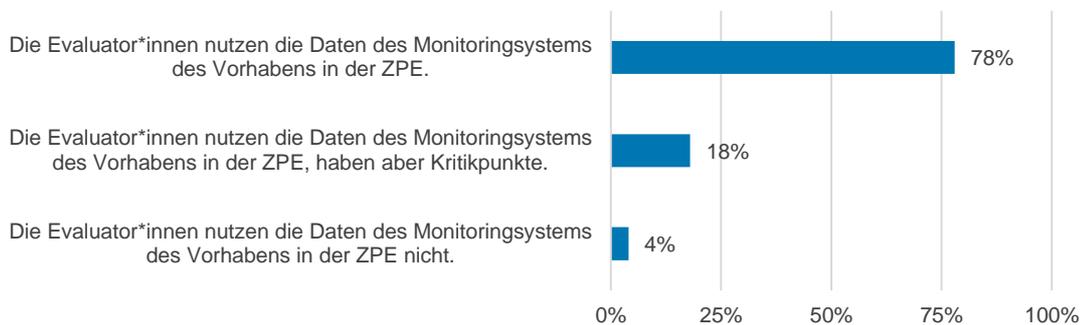
<sup>9</sup> Aus der Stichprobe (50 Vorhaben) gehörten dabei 27 Vorhaben einem Programm der internationalen Zusammenarbeit an.

## Bewertung des WoM

Wie bereits mehrfach erwähnt, sollen die Monitoringdaten des Vorhabens die Evaluationsteams in der ZPE unterstützen. In der QSA wurde daher untersucht, inwieweit die Monitoringdaten von den Evaluationsteams tatsächlich genutzt wurden. Dabei wird deutlich, dass die Evaluationsteams die Daten aus den Monitoringsystemen der Vorhaben insbesondere zur Validierung von Wirkungshypothesen verwenden. Die Ergebnisse der Untersuchung, inwiefern die Monitoringdaten zur ZPE genutzt werden, sind in Abbildung 29 dargestellt. Bei genauerer Betrachtung wird deutlich, dass die Evaluatoren\*innen die Daten in neun von zehn Vorhaben mindestens teilweise nutzen, um die Wirkungshypothesen auf der Outcome-Ebene zu überprüfen (siehe Abbildung 30). In einem etwas geringeren Maß, nämlich in etwa sieben von zehn Vorhaben, nutzen sie die Monitoringdaten, um Wirkungshypothesen auf der Impact-Ebene zu überprüfen und zu validieren (siehe Abbildung 31). Da die erhobenen Daten für die Evaluierung genutzt werden, ist die Qualität des WoM hinsichtlich der Evaluierbarkeit damit als gut zu bewerten.

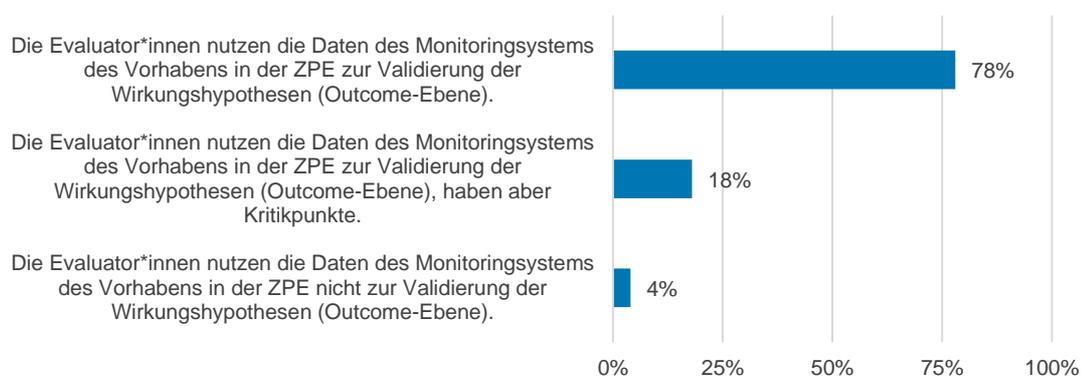
Außerdem wurden die Teilnehmer\*innen der Deep Dives gebeten, eine Hypothese zur Synchronisierung zwischen Monitoring und Evaluierung zu bewerten. Die Hypothese lautet: „**Wenn das Monitoringsystem gut ist, werden Monitoringdaten in der Evaluierung des Vorhabens genutzt.**“ (Hypothese 5). Diese Hypothese wurde von allen fünf befragten Projektteams **bestätigt**. Sie hoben einerseits hervor, dass das jeweilige Evaluierungsteam der ZPE darauf hingewiesen habe, wie hilfreich die Monitoringdaten waren. Die Interviewten betonten, dass andererseits berücksichtigt werden müsse, dass die Monitoringdaten auch in der ZPE genutzt würden, falls die Monitoringdaten nicht gut seien.

Abbildung 22: Nutzung der Monitoringdaten in den ZPE



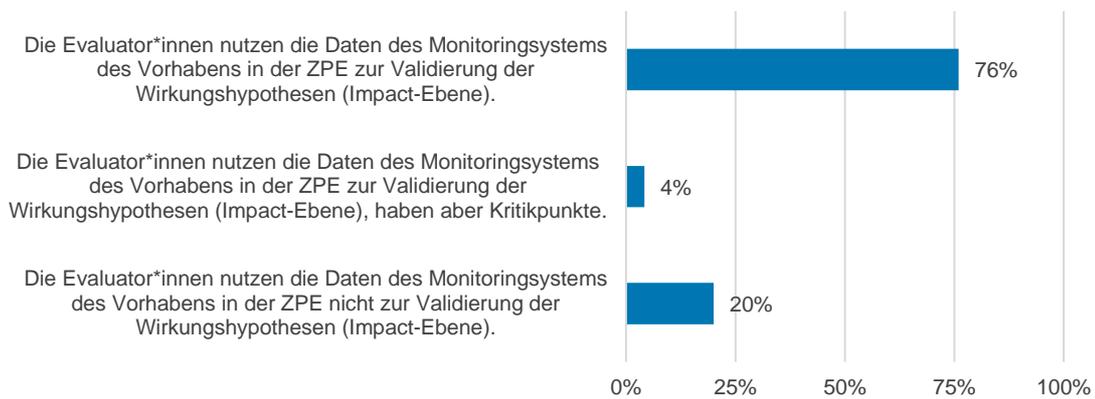
Quelle: Syspons 2022

Abbildung 23: Nutzung der Monitoringdaten in den ZPE zur Validierung der Wirkungshypothesen (Outcome-Ebene)



Quelle: Syspons 2022

Abbildung 24: Nutzung der Monitoringdaten in den ZPE zur Validierung der Wirkungshypothesen (Impact-Ebene)



Quelle: Syspons 2022

### Evaluability Assessment

Die Stabsstelle Evaluierung hat ein *Evaluability Assessment* erstellt, das heißt eine Checkliste, um die Evaluierbarkeit einer Maßnahme zu prüfen. Die QSA kommt zu dem Schluss, dass das *Evaluability Assessment* durch die aufgeworfenen Fragestellungen zur Skalierbarkeit der Erkenntnisse der Evaluierungen beitragen kann. Das *Evaluability Assessment* ist ein detailliertes Formular zur selbstständigen Prüfung durch Vorhaben, ob diese insbesondere für dezentrale Projektevaluierungen geeignet sind. In der Querschnittsauswertung gab es jedoch keine konkreten Hinweise darauf, inwiefern das *Evaluability Assessment* der STS Evaluierung genutzt werden kann, da dieses unter den Projektteams nicht bekannt war. Im Rahmen der Analyse des *Evaluability Assessment* wurde jedoch deutlich, dass es, in Abhängigkeit davon, wofür es verwendet werden soll, potenziell von Interesse ist und eine Schnittstellenfunktion zwischen Monitoring und Evaluierung einnehmen kann. So wirft das *Evaluability Assessment* Fragen wie den realistischen Nutzen und/oder die Skalierbarkeit der Erkenntnisse einer Evaluierung auf (siehe Abbildung 32). Wird es zum Beispiel im Rahmen des Monitorings der Vorhaben eingesetzt und auf dieses angepasst, kann es durch die damit angestoßene Reflexion dazu beitragen, dass die Erkenntnisse der Erhebungen skalierbar und damit für GIZ-interne Akteure zur Planung von Vorhaben von Interesse sind. Damit kann das *Evaluability Assessment* mindestens zu den Dimensionen Monitoring-Evaluierung sowie Evaluierung-Planung beitragen.

Abbildung 25: *Evaluability Assessment* (Ausschnitt) (Quelle: GIZ).

Use	
Guiding question	Remarks
What is the main purpose of the evaluation?	
Is the desired use realistic?	
Is the timing right?	
Is there an opportunity for an evaluation to have an influence?	
Has the project accumulated enough implementation experience to enable useful lessons to be extracted?	
Are there mechanisms or strategies in place to strengthen the use of evaluation findings (e.g. networks, targeted events or other platforms to discuss and share findings)?	
Are findings scalable?	
Do they increase visibility of the project?	

## Erfolgs- und Misserfolgskfaktoren

**Die bisherigen Vorgaben/Handreichungen sowie das gemeinsame Verständnis im Projektteam, dass Monitoring relevant für die Projektsteuerung ist, gelten als Erfolgsfaktoren für eine Synchronisierung zwischen Monitoring und Evaluierung.** So wurde in den Gesprächen mit den Projektteams oft auf die Vorgaben/Handreichungen zur Erstellung der Indikatoren hingewiesen, auf deren Basis während der Planung der Vorhaben SMARTe Indikatoren erstellt worden seien (siehe Kapitel 3.1). Obgleich die Indikatoren in vielen ZPE teilweise als nicht-SMART bewertet wurden (dies betraf besonders die Dimension der Spezifität), wurde damit teilweise die Qualität der Indikatoren, die für die ZPE genutzt werden können, sichergestellt. Im Rahmen der Deep Dives wurde festgestellt, dass ein weiterer Erfolgsfaktor ist, ob in den (interviewten) Projektteams ein gemeinsames Verständnis darüber herrscht, dass Monitoring eine relevante Rolle in der Steuerung der Vorhaben spielt. Durch die Bedeutung des Monitorings in den Entscheidungen des Vorhabens gewinnt es als Grundlage für die Evaluierung des Vorhabens an Relevanz.

**In der Auswertung hat sich herausgestellt, dass die für das Monitoring bereitgestellten Ressourcen sowie das lückenhafte Verständnis in Bezug auf konfliktsensibles Monitoring Misserfolgskfaktoren für eine Synchronisierung zwischen Monitoring und Evaluierung sind.** Wie im vorherigen Punkt zur Verzahnung zwischen Planung und Monitoring, ist auch für die Verzahnung zwischen Monitoring und Evaluierung relevant, dass im Vorhaben angemessene Ressourcen für ein erfolgreiches Monitoring eingeplant werden (siehe Kapitel 3.1). Es stellte sich als ein weiterer Misserfolgskfaktor heraus, der eine erfolgreiche Nutzung des Monitoringsystems in der Evaluierung erschwert, dass in den Projektteams ein gemeinsames beziehungsweise korrektes Verständnis für konfliktsensibles Monitoring fehlt. Dieses wurde in den Gesprächen oft mit dem Risikomonitoring der Vorhaben gleichgesetzt.

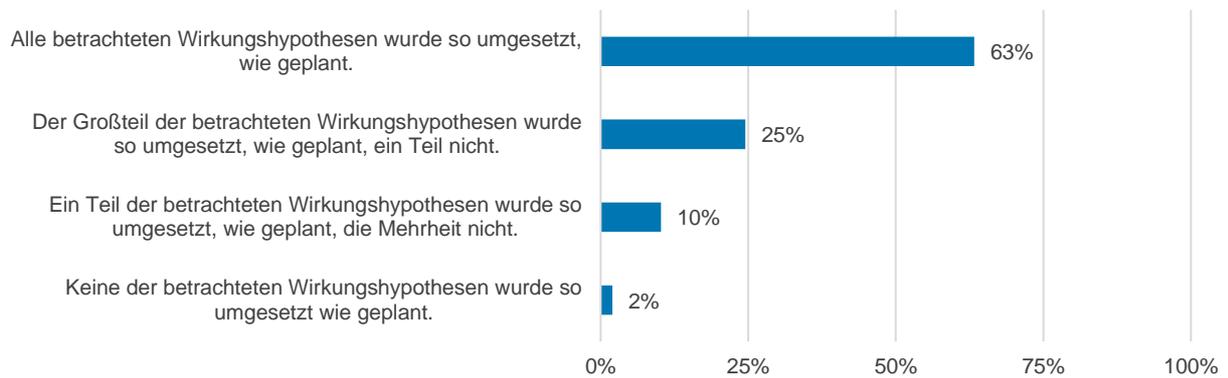
## 3.3 Evaluierung - Planung

Eine **Zielsetzung hinter den Evaluierungen der STS Evaluierung ist, dass diese die Planung von Vorhaben der GIZ unterstützen.** So haben ZPE von durch das BMZ beauftragten Vorhaben drei Grundfunktionen: Unterstützung evidenzbasierter Entscheidungen, Transparenz und Rechenschaftslegung sowie organisationales Lernen im Sinne eines Beitrags zu effektivem Wissensmanagement. Die Querschnittsauswertung hat daher die Frage untersucht, welche Potenziale (ZPE-)Evaluierungsergebnisse auf den verschiedenen Ebenen innerhalb der GIZ für die Planung von Neu- und/oder Folgevorhaben sowie für das Lernen (im Land, Sektor und/oder Organisationales Lernen) bieten, und wie sie genutzt werden. Dabei werden zuerst die Erkenntnisse zur Implementierung der Wirkungshypothesen aufgeführt, gefolgt von Erkenntnissen zur Verzahnung mit Vorgängervorhaben und zur Nutzung von Evidenzen für Folgevorhaben.

### Implementierung der Wirkungshypothesen

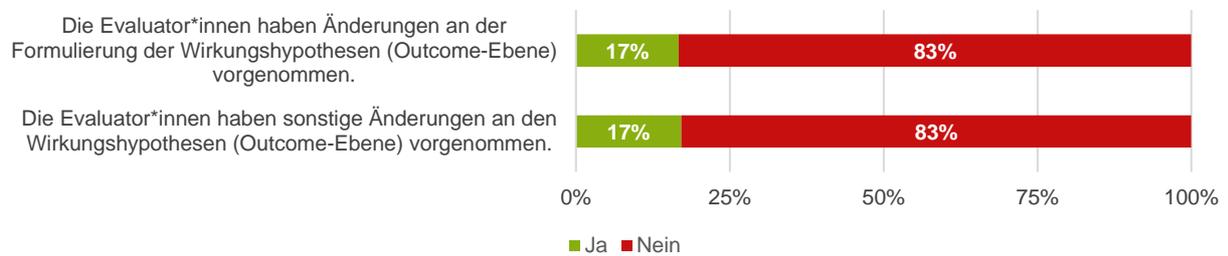
**Damit die Ergebnisse der Evaluierung in der Planung neuer Vorhaben genutzt werden können, ist es zentral, für die Planung relevante Aspekte wie die Wirkungshypothesen auf den verschiedenen Ebenen zu betrachten.** Die Auswahl und Bewertung von drei Wirkungshypothesen auf der Outcome-Ebene und zwei Wirkungshypothesen auf der Impact-Ebene ist dabei Standard im ZPE-Prozess. Die ausgewählten Wirkungshypothesen werden in der Evaluierung von Evaluationsteams entweder bestätigt, teilweise bestätigt oder nicht bestätigt. **Die qualitative Auswertung zeigt dabei, dass der Großteil der angepassten Wirkungshypothesen auf der Outcome-Ebene von Evaluationsteams bestätigt wurde und damit der Realität der Implementierung entspricht.** Die Ergebnisse dieser Auswertung sind in Abbildung 33 dargestellt. Dabei gibt es in nur ungefähr zwei von zehn ZPE-Berichten Hinweise darauf, dass die Evaluationsteams Änderungen an den Wirkungshypothesen auf der Outcome-Ebene vorgenommen haben (siehe Abbildung 34). Da der Großteil der Wirkungshypothesen durch die Evaluationsteams bestätigt wurde, eignen sie sich damit für die Planung weiterer Vorhaben.

Abbildung 26: Umsetzung der Wirkungshypothesen (Outcome-Ebene) der untersuchten Vorhaben



Quelle: Syspons 2022

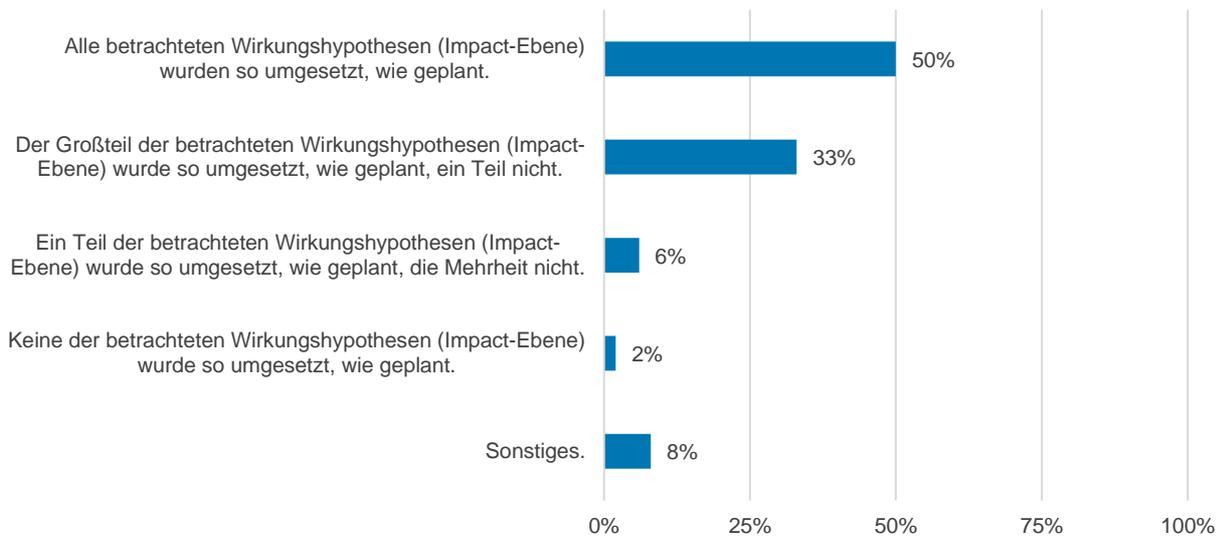
Abbildung 27: Änderungen an Wirkungshypothesen (Outcome-Ebene) der untersuchten Vorhaben in der ZPE



Quelle: Syspons 2022

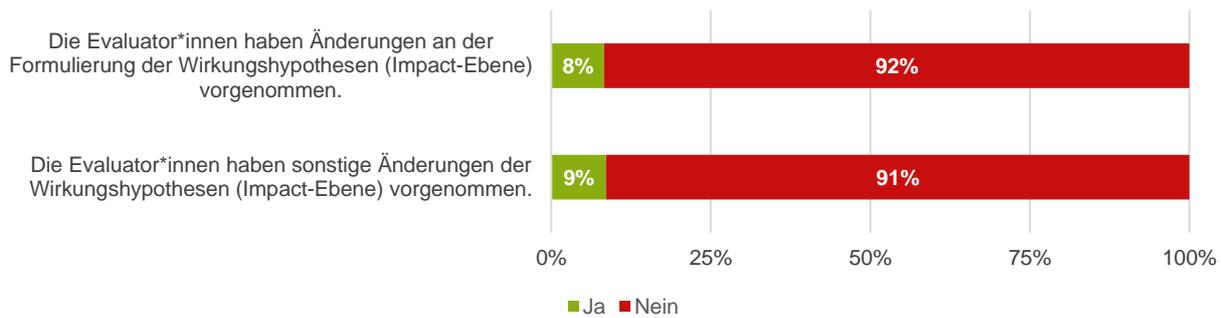
**Auch der Großteil der angepassten Wirkungshypothesen auf der Impact-Ebene wurde von Evaluationsteams bestätigt und entspricht damit der Realität der Implementierung.** Die Ergebnisse dieser Auswertung sind in Abbildung 35 dargestellt. In nur einem von zehn ZPE-Berichten gibt es Hinweise darauf, dass die Evaluationsteams Änderungen an den Wirkungshypothesen auf der Impact-Ebene vorgenommen haben (siehe Abbildung 36). Damit eignen sich auch die Wirkungshypothesen auf der Impact-Ebene für die Planung weiterer Vorhaben.

Abbildung 28: Umsetzung der Wirkungshypothesen (Impact-Ebene) der untersuchten Vorhaben



Quelle: Syspons 2022

Abbildung 29: Änderungen an Wirkungshypothesen (Impact-Ebene) der untersuchten Vorhaben in der ZPE

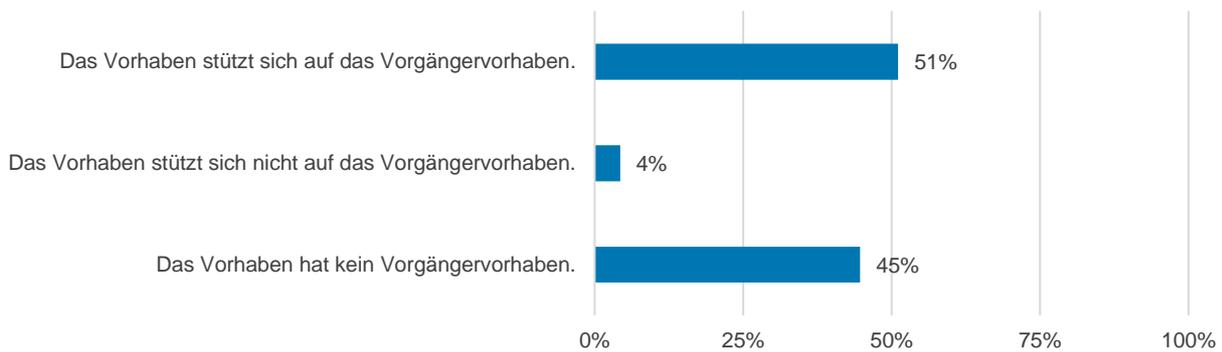


Quelle: Syspons 2022

### Vorgängervorhaben

Ein zentraler Zweck der ZPE ist, dass ihre Erkenntnisse in die Planung und Umsetzung der Folgevorhaben einfließen. Daher wurden von den Auswerter\*innen die in den ZPE-Berichten enthaltene Informationen erfasst, um die Erfahrungen der Vorgängervorhaben nutzen zu können. In den ZPE-Berichten ist dabei deutlich zu erkennen, dass sich fast alle Vorhaben mit Vorgängervorhaben auf die Erfahrungen ihrer Vorgänger (zum Beispiel durch persönliches Wissen der Projektbeteiligten) stützen. Die Ergebnisse dieser Frage sind in Abbildung 37 dargestellt. So werden zum Beispiel Handlungsfelder und/oder Partnerstrukturen leicht modifiziert beibehalten. Ein Beispiel für gute Praktiken für die Nutzung der Erkenntnisse des Vorgängervorhabens ist in Abbildung 38 dargestellt. Gleichzeitig war zu erkennen, dass es in den ZPE-Berichten nur selten Hinweise darauf gab, inwiefern Erkenntnisse / Empfehlungen aus Evaluierungen der Vorgängervorhaben umgesetzt werden. Dies kann jedoch auch damit zusammenhängen, dass das Vorgängervorhaben nicht evaluiert wurde und es damit keine expliziten externen Erkenntnisse gibt, die hätten genutzt werden können.

Abbildung 30: Nutzung der Erfahrungen der Vorgängervorhaben



Quelle: Syspons 2022

Kasten 8: Beispiel für gute Praktiken: Nutzung der Erfahrungen der Vorgängervorhaben

## Beispiel für gute Praktiken

- **Politikberatung im Umwelt- und Klimaschutz (PAKLIM)** (PN 2015.2117.8): Im ZPE-Bericht wird darauf hingewiesen, dass das Vorgängervorhaben evaluiert und ein Kritikpunkt umgesetzt wurde. In der ZPE wurde der Fokus des Vorhabens auf die Kapazitätsentwicklung einzelner Akteure kritisiert. Im betrachteten Vorhaben wurde stattdessen in den Kapazitätsaufbau auf Organisationsebene investiert.

### Folgevorhaben

Außerdem war ein zentraler Aspekt der QSA, zu betrachten, inwiefern Ergebnisse der ZPE der Stichprobe der QSA PME in den Folgevorhaben der evaluierten Vorhaben genutzt wurden. Aus den Rückmeldungen der AVs sowie den Gesprächen im Rahmen der Deep Dives wird sichtbar, dass Erfahrungen aus der Evaluierung zum Teil für Folgevorhaben genutzt wurden. Beispiele sind in anonymisierter Form in Abbildung 39 aufgeführt.

Kasten 9: Ausgewählte Rückmeldungen der AV-(V-)Befragung: Nutzung der ZPE-Ergebnisse im Folgevorhaben

## Ausgewählte Rückmeldungen der AV-Befragung (anonymisiert)

- **Rückmeldung 1:** „Ich habe für die [neue] Laufzeit des Vorhabens ein formales Risikomanagementsystem eingeführt. Das bestand zwar schon vorher, allerdings [auf einer anderen Ebene], und nicht vollständig formalisiert. Dies hat die ZPE-Gutachter zu einer tieferen Punktvergabe veranlasst.“
- **Rückmeldung 2:** „In der neuen Phase wurden z.T. aufgrund der Empfehlungen neue Strukturen im Vorhaben geschaffen, M&E-Prozesse wurden verbessert, um auch „on the go“ Impacts festzuhalten.“ [...].“

**Dennoch bildet die zentrale Datenquelle für die Planung von Vorhaben der eigene Erfahrungsschatz der involvierten Akteur\*innen.** Aus der Rückmeldung der explorativen Interviews unter anderem mit Fachplaner\*innen sowie aus den (Gruppen-) Gesprächen im Rahmen der Deep Dives wird sichtbar, dass der eigene Erfahrungsschatz der in die Planung involvierten Akteur\*innen (Fachplaner\*innen, AV-Vs) die zentrale Datenquelle für die Planung der Neu-/Folgevorhaben darstellt. **Die Dokumentenauswertung zeigt zudem, dass die Nutzung von Evidenzen nicht vorgeschrieben ist.** In den betrachteten Handreichungen sowie weiteren Unterlagen wie Verfahrensinformationen gibt es keine Hinweise und/oder Vorgaben zur Nutzung von Evidenzen in der Planung von Vorhaben. Die einzige Ausnahme, die allerdings nur bedingt auf Vorhaben einwirkt, ist die Verfahrensinformation des BMZ zur Erstellung von Länderstrategien (GIZ 2021). In diesem Sinn wird deutlich, dass der Bewertung von Wirkungshypothesen durch ZPE eher kein Gewicht beigemessen wird. Beispiele sind in anonymisierter Form in Abbildung 40 aufgeführt.

Kasten 10: Ausgewählte Rückmeldungen der AV-(V)Befragung: Nutzung des eigenen Erfahrungsschatzes

## Ausgewählte Rückmeldung der Deep Dives (anonymisiert)

- **Rückmeldung 1:** „Das waren alle Bereiche, in denen wir schon tätig waren. Da konnten wir sehr gut aufgrund unserer eigenen Erfahrung einschätzen. [...] Da beruhte die Konzeption sehr auf den Erfahrungen, die wir bereits gemacht hatten. Stark evidenzbasiert, nicht aufgrund einer Studie oder ähnlichem, sondern der Erfahrung aus der vorhergehenden Kooperation.“
- **Rückmeldung 2:** „Wir haben zur Planung Erfahrung aus der Zusammenarbeit mit den gleichen Partnern geschöpft und uns darauf berufen. [...] Wissenschaftliche Literatur haben wir nicht herangezogen, sondern GIZ-interne Vorgaben [...]. Und Abstimmung mit Partnern im Sektor. [Aus meinen vorherigen beruflichen Tätigkeiten] habe ich [...] Kenntnisse mitbekommen.“

**Der Anspruch der GIZ ist es, dass sich Planung, Monitoring und Evaluierung der Vorhaben miteinander verzahnen und sich gegenseitig unterstützen. So sollen Wirkungsmodell und Indikatoren die Etablierung eines angemessenen WoMs erleichtern, die etablierten Monitoringsysteme eine geeignete Grundlage für spätere Evaluierungen darstellen, und Evaluierungsergebnisse innerhalb der GIZ für die Planung von Folge- und/oder Neuvorhaben sowie sektorales und organisationales Lernen genutzt werden.**

Um diesen Zusammenhang zu überprüfen, wurden die Teilnehmer\*innen der Deep Dives um ihre Einschätzung einer entsprechenden Hypothese gebeten. Die Hypothese lautet: *„Wenn das Monitoringsystem gut ist, unterstützt dies eine gute Evaluierung, deren Erkenntnisse in die Planung neuer und/oder Folgevorhaben einfließen können.“* (**Hypothese 6**). Diese Hypothese wurde von allen befragten Projektteams **bestätigt**. Sie wiesen jedoch auf mehrere Hindernisse einer erfolgreichen Synchronisierung von Planung, Monitoring und Evaluierung hin. Einerseits muss der Faktor ‚Mensch‘ berücksichtigt werden, der sich, zum Beispiel in der Form von Fachkenntnissen und/oder persönlichem Engagement, auf die Anwendung der Monitoringsysteme auswirken kann. Andererseits ist der zeitliche Zusammenhang zwischen ZPE und der Planung von Folgevorhaben eine wesentliche Einschränkung für die Nutzung der gewonnenen Evidenzen: So werden Folgevorhaben meist vor der Bereitstellung der ZPE-Ergebnisse geplant, geprüft und in Auftrag gegeben.

## Erfolgs- und Misserfolgskfaktoren

**In der Auswertung hat sich das rechtzeitige Bereitstellen von relevanter Information als Erfolgsfaktor für eine Synchronisierung zwischen Evaluierung und Planung herausgestellt.** Um Evidenzen in der Planung von Neu- und/oder Folgevorhaben nutzen zu können, müssen diese denjenigen, die in die Planungs- und gegebenenfalls in die Prüfprozesse involviert sind, zur Verfügung stehen. Dies gilt insbesondere für Evidenzen aus den ZPE der Vorgängervorhaben. Diese konnten in der Vergangenheit, laut den Interviewten, insbesondere dadurch genutzt werden, dass die AVs/AV-Vs bereits im ZPE-Prozess involviert waren und daher Ergebnisse bereits vor der Veröffentlichung der ZPE-Berichte kannten und somit in der Planung nutzen konnten.

**In der Auswertung hat sich der zeitliche Rahmen der ZPE und Planung von (Folge-)Vorhaben sowie eine fehlende Vergemeinschaftung von Evaluierungsergebnissen als Misserfolgskfaktor für eine Synchronisierung zwischen Evaluierung und Planung herausgestellt.** Wie bereits aufgeführt, finden ZPE oft erst parallel zur oder nach der Planung der Folgevorhaben statt, und ihre Ergebnisse können oft erst in der Umsetzung der Folgevorhaben genutzt werden. Mit Blick auf weitere Vorhaben fehlt es an einer Vergemeinschaftung von Evaluierungsergebnissen, zum Beispiel mit anderen GIZ-Vorhaben im gleichen Partnerland/-region oder im gleichen Sektor. Dies schränkt die Nutzung von Evaluierungsergebnissen für ähnliche Vorhaben weiter ein.

## 4 Schlussfolgerungen und Empfehlungen

Aufgrund der **dargestellten Erkenntnisse der qualitativen Untersuchung lassen sich die folgenden Schlussfolgerungen ziehen.**

**Die QSA zeigt, dass die GIZ in der Planung in bestimmten Bereichen gut aufgestellt ist.** So verfügen die Vorhaben zu Beginn der ZPE überwiegend über eigene systemische Wirkungsmodelle, die Impacts, Outcomes und Outputs miteinschließen, und über eindeutig ersichtliche und angemessene Systemgrenzen. Allerdings ist aus den ZPE-Berichten nicht ersichtbar, inwiefern die Wirkungsmodelle während der Planung oder Umsetzung der Vorhaben erstellt wurden. Diese **Wirkungsmodelle** unterstützen die Projektteams im (späteren) Monitoring des Vorhabens, da sie insbesondere zu einem gemeinsamen Verständnis der Ziele und Wirkmechanismen beitragen.

Allerdings zeigt die qualitative Auswertung, dass das Wirkungsmodell während der ZPE überwiegend vom Evaluationsteam angepasst wird, um der Realität des Vorhabens zu entsprechen und eine geeignete Grundlage für die Evaluierung zu bilden. Übergeordnete Aussagen zur Qualität der Zielformulierungen und der Beziehung zwischen Zielen und Indikatoren sind nicht möglich, da diese von den Evaluationsteams nur selten explizit betrachtet wurden. Mit Blick auf die mit dem Auftraggeber vereinbarten **Indikatoren** selbst wurde deutlich, dass die Erstellung von SMARTen Indikatoren verbindlich ist und in der ZPE auf der Outcome-Ebene geprüft wird. Durch die verbindliche Vorgabe zur Erstellung SMARTer Indikatoren soll die Angemessenheit der Indikatoren sichergestellt werden, um ein effektives WoM zu etablieren. Es zeigt sich aber, dass die Evaluationsteams in der Hälfte der ZPE mindestens einen Teil der geprüften Indikatoren kritisieren und insbesondere ihre Spezifität und Messbarkeit anpassen. Ziel der Anpassungen ist es, dass die Indikatoren von den Evaluationsteams für die Evaluierung der Vorhaben genutzt werden können. Insgesamt wird die Qualität der Indikatorensysteme von den Evaluationsteams allerdings als gut bewertet. So sind die vereinbarten Baseline- und Zielwerte überwiegend nachvollziehbar und eignen sich somit für das fortlaufende Monitoring der Vorhaben. Allerdings wurde deutlich, dass es noch Potenzial bei der Nutzung von **Datenquellen** wie Partnersystemen

sowie öffentlichen Statistiken gibt. Als einschränkend wird zudem die Relevanz von angemessenen geplanten **Personalressourcen** für adäquates Monitoring aufgeführt. In diesem Sinne war ersichtlich, dass in der Verknüpfung zwischen Planung und Monitoring noch Potenziale bestehen.

**Die QSA zeigt weiterhin, dass die vorliegenden WoM-Systeme eine spätere Evaluierung der Vorhaben unterstützen.** So gibt es in den Verfahrensinformationen des BMZ und den Handreichungen der GIZ explizite Hinweise und Vorgaben, die das Monitoring der Vorhaben betreffen und zum Beispiel den Aufbau des Monitoringsystems und/oder praktische Aspekte der Datenerhebung thematisieren. Bei der qualitativen Betrachtung der ZPE-Berichte und in den Deep Dives wurde deutlich, dass in mehr als der Hälfte der Vorhaben regelmäßig selbstständig **Monitoringenerhebungen** durchgeführt werden und Monitoringdaten zur **Steuerung** der Vorhaben genutzt werden. Allerdings wird seltener von **Monitoringmeetings** berichtet. Zudem werden Datenerhebungsmethoden wie die offene Perspektivenerfassung KOMPASS, die gezielt Partnerorganisationen und Zielgruppen einbinden, bislang von eher wenigen Vorhaben genutzt. Auch werden **nicht-intendierte Wirkungen** bislang nur selten erhoben und es gibt Informationsbedarf bezüglich eines **konfliktsensiblen Monitorings**. Letzteres wird von Projektteams oft mit Risikomanagement gleichgesetzt. Dennoch war ersichtlich, dass die Evaluationsteams stets die Daten des Monitoringsystems der Vorhaben zur ZPE nutzen, insbesondere zur **Überprüfung der Wirkungshypothesen** auf der Outcome- sowie der Impact-Ebene<sup>10</sup>. Unabhängig von den ZPE wollte die GIZ die Vorhaben durch das Instrument des **Evaluability Assessments**<sup>11</sup> in der Durchführung dezentraler Projektevaluierungen unterstützen. Die Querschnittauswertung hat allerdings gezeigt, dass das *Evaluability Assessment* zwar Potenziale für eine weitere Verzahnung zwischen Monitoring und Evaluierung aufweist, in den Projektteams bislang allerdings noch nicht bekannt ist.

**Zwar werden Erkenntnisse aus den ZPE in (Folge-)Vorhaben genutzt, allerdings sind die eigenen Erfahrungswerte der Beteiligten die zentrale Evidenz.** Bei der qualitativen Auswertung sowie den Deep Dives wurde deutlich, dass die Nutzung von Erkenntnissen der ZPE darin begründet ist, dass insbesondere AVs der evaluierten Vorhaben an der Planung der (Folge-)Vorhaben beteiligt sind und die Erkenntnisse der ZPE als Teil ihres eigenen Erfahrungsschatzes in die Planung einbringen. Die Nutzung von Evidenzen ist in der GIZ jedoch **nicht systematisch verankert, nicht vorgegeben**, und wird zudem durch den **zeitlichen Rahmen** der ZPE und der Planung der (Folge-)Vorhaben erschwert. Dabei zeigt die qualitative Auswertung, dass sich gerade die **Wirkungshypothesen** der Vorhaben für die Planung von (Folge-)Vorhaben eignen, da die bestätigten Wirkungshypothesen eine zentrale Grundlage für eine weitere evidenzbasierte Planung liefern.

**Auf der Basis der konkreten Erkenntnisse und Schlussfolgerungen aus dieser Evaluierung spricht das Evaluierungsteam folgende Empfehlungen aus:**

---

<sup>10</sup> Dabei muss darauf hingewiesen werden, dass die Überprüfung der Wirkungshypothesen indirekt und nicht direkt auf den Monitoringdaten basiert. Bei der Überprüfung der Wirkungshypothesen wird unter anderem geprüft, inwiefern diese die Realität des Vorhabens abbilden. Dabei wird ebenfalls geprüft, inwiefern die Ziele (insbesondere auf der Outcome-Ebene) erreicht wurden. In dieser Bewertung wird auch die Zielerreichung der Indikatoren untersucht.

<sup>11</sup> Ein von der GIZ entwickeltes Instrument: Ein detailliertes Formular zur selbstständigen Prüfung, ob das eigene Vorhaben für eine dezentrale Evaluierung geeignet ist.

## Übergeordnete Handlungsempfehlung

**Die Auswertung hat gezeigt, dass bislang ein Verständnis für die Integration von konfliktensiblen Indikatoren fehlt, beziehungsweise Projektteams dieses losgelöst von ihrem regulären Monitoring sowie Risikomonitoring betrachten. Die GIZ versteht Monitoring, Risikomonitoring und konfliktensibles Monitoring allerdings als integratives Konzept. Daher wird der GIZ empfohlen, eine Informationskampagne zu konfliktensiblen Monitoring zu starten.** In der Auswertung wurde deutlich, dass die Projektteams zwar gewillt waren, Informationen zu Risiken für das Projekt zu erheben und zu besprechen; allerdings war ebenfalls zu erkennen, dass den Projektteams Informationen und Kenntnisse zu konfliktensiblen Monitoring fehlten und dass konfliktensiblen Indikatoren von den Vorhaben nur selten erfasst wurden. Das liegt insbesondere darin begründet, dass die Projektteams Monitoring, Risikomonitoring und konfliktensibles Monitoring als voneinander losgelöst betrachten. So wurde mit Blick auf die Erfassung von Risiken auch mehrfach auf das Risikomonitoring durch die Risk Management Officer der Landesbüros verwiesen, die allerdings Risiken auf Ebene des Partnerlandes und nicht projektspezifische Informationen erheben. Ein integratives Verständnis ist essenziell, da in den Monitoringansätzen unterschiedliche Aspekte betrachtet werden, die alle relevant für das Vorhaben sind: Während beim Risikomonitoring Risiken für die erfolgreiche Umsetzung des Vorhabens erfasst werden, stehen im konfliktensiblen Monitoring (insbesondere konfliktfördernde und/oder -mindernde) Wirkungen des Vorhabens auf den Kontext des Partnerlandes im Vordergrund. Dadurch wird die Betrachtung nichtintendierter Wirkungen gefördert, die bislang im Monitoring der Vorhaben eine untergeordnete Rolle spielen. Die GIZ sollte daher relevante Mitarbeiter\*innen der Projekte (AVs, Monitoringbeauftragte) sowie Fachplaner\*innen zu konfliktensiblen Monitoring im Rahmen einer Informationskampagne weiterbilden. Diese Informationskampagne hat das Ziel, dazu beizutragen, konfliktensibles Monitoring in den Projektteams bekannter zu machen. Dabei sollte konfliktensibles Monitoring auch weiterhin als Bestandteil eines integrativen Monitoringkonzepts vermittelt werden. Damit diese Indikatoren vermehrt Anwendung finden, sollte sie zudem durch weiterführendes Material (zum Beispiel Handreichungen) begleitet werden.

## Handlungsempfehlung: Synchronisierung von Planung und Monitoring

**Den Vorhaben wird empfohlen, im Personalkonzept ausreichend Personalressourcen für Monitoring bereitzustellen.** Die Auswertung hat gezeigt, dass ein zentraler Erfolgsfaktor für eine erfolgreiche Synchronisierung von Planung und Monitoring beziehungsweise Monitoring und Evaluierung ist, dass angemessene Ressourcen zur Konzeption, Umsetzung und Auswertung (und gegebenenfalls Anpassung) eines Monitoringsystems vorhanden sind. Daher ist es wichtig und wird den Vorhaben empfohlen, ausreichend Personalressourcen für das Monitoringsystem, zum Beispiel in der Form von Monitoringbeauftragten mit entsprechendem Mandat und entsprechender Rolle im Vorhaben, einzuplanen. Es geht allerdings weniger darum, mehr Ressourcen bereitzustellen, sondern darum, mehr Ressourcen an der richtigen Stelle zu fokussieren. So können durch mehr Ressourcen für Monitoring anderweitig Ressourcen eingespart werden (zum Beispiel bei der Operationsplanung), wenn Monitoringdaten auch für die strategische Steuerung der Vorhaben genutzt werden. Dagegen sollten, wo angemessen, auch weitere Projektmitarbeiter\*innen in das Monitoring einbezogen werden: zum Beispiel dadurch, dass die Zulieferung von Daten (an die Monitoringbeauftragten) in ihren Rollenbeschreibungen verankert wird.

**Die durch Qualitätssicherung in Linie (QSIL)<sup>12</sup> bestehende Verpflichtung für jedes Vorhaben, ein Wirkungsmodell und WoM zu erstellen, sollte beibehalten und nachgehalten werden, da sich dies vorteilhaft auf die Qualität der Planung und Evaluierung auswirkt.** Die Auswertung hat gezeigt, dass die von den Vorhaben erstellten Wirkungsmodelle – gegebenenfalls von den Evaluationsteams angepasst – sich positiv auf die Vorhaben auswirken, da sie ein gemeinsames Verständnis der Ziele und Wirkmechanismen im Projektteam ermöglichen. Zudem ermöglichen sie im Rahmen der ZPE, die Ziele und Wirkmechanismen der Vorhaben zu

<sup>12</sup> Qualitätssicherung in Linie (QSIL) der GIZ umfasst die zwischen Auftragsverantwortlichen und Führungskraft schriftliche vereinbarten Mindeststandards (Leistungserbringung, Führungs- und Qualitätsdialog) für die inhaltlich-fachliche und die kaufmännische Qualität in der Durchführung von Aufträgen des Gemeinnützigen Bereichs.

begreifen und ihren Erfolg zu untersuchen. Ähnliches gilt für die WoM-Systeme, die es Projekt- und später Evaluationsteams ermöglichen, die Wirkungen des Vorhabens zu beobachten und zu erfassen. Wichtig ist dabei, dass diese Dokumente nachgehalten, das heißt regelmäßig vom Projektteam reflektiert und gegebenenfalls aktualisiert werden. Die WoM sollten weiterhin, begleitend zu den oben sowie nachfolgend genannten Empfehlungen, um konfliktsensible Indikatoren sowie um die Erfassung von Informationen zu Partnerorganisationen und Zielgruppen ergänzt werden.

### **Handlungsempfehlung: Synchronisierung von Monitoring und Evaluierung**

**Auch mit Blick auf die Synchronisierung von Monitoring und Evaluierung wird den Vorhaben empfohlen, im Personalkonzept ausreichend Personalressourcen für Monitoring bereitzustellen.** Dabei muss das laufende Monitoring, und, falls gezogen, auch die ZPE mitbedacht werden.

**Zudem wird den Vorhaben empfohlen, nicht-intendierte Wirkungen sowie die Perspektive der Zielgruppe stärker zu erfassen.** Die Auswertung hat deutlich gemacht, dass Instrumente zur Erhebung weiterer Informationen zu nicht-intendierten Wirkungen sowie den Partnerorganisationen und Zielgruppen teilweise eher selten genutzt werden. Daher wird den Vorhaben empfohlen, dies stärker in ihr Monitoring einzubauen, zum Beispiel durch die Nutzung der Instrumente der offenen Perspektivenerfassung KOMPASS. Die bisherige Perspektivenerfassung der Partnerorganisationen und/oder Zielgruppen, die zum Beispiel auf nicht-intendierte Wirkungen hinweisen kann, wird aktuell meist nur im Rahmen der Erhebung der abgestimmten Indikatoren genutzt.

**Außerdem wird der GIZ empfohlen, die Einsatzmöglichkeiten des *Evaluability Assessments* auf Ebene der Vorhaben zu prüfen und das *Evaluability Assessment* gegebenenfalls entsprechend anzupassen.** In der Querschnittauswertung wurde deutlich, dass das von der GIZ entwickelte *Evaluability Assessment* – ein detailliertes Formular zur selbstständigen Prüfung, ob das eigene Vorhaben für eine dezentrale Evaluierung geeignet ist – in den Projektteams noch nicht bekannt ist. In der genaueren Betrachtung wurde jedoch auch deutlich, dass das *Evaluability Assessment* unterschiedliche Potenziale und Einsatzmöglichkeiten für die Projektteams aufweist. So kann das *Evaluability Assessment* erstens eine dezentrale Evaluierung des Vorhabens vorbereiten. Zweitens kann das *Evaluability Assessment* von den Vorhaben genutzt werden, um zu prüfen, wie sie in ihrem Monitoring aufgestellt sind. Eine dritte potenzielle Funktion besteht darin, dass das *Evaluability Assessment* von Vorhaben genutzt werden kann, um ihre eigenen Studien (zum Beispiel Baselinestudien) zu prüfen und zu planen. Daher sollten die Einsatzmöglichkeiten des *Evaluability Assessment* auf der Ebene der Vorhaben geprüft und das Instrument gegebenenfalls an die unterschiedlichen Funktionen, die es erfüllen kann, inhaltlich angepasst werden (zum Beispiel Filter spezifischer Fragen je nach Einsatzmöglichkeit).

### **Handlungsempfehlung: Synchronisierung von Evaluierung und Planung**

**Der GIZ wird empfohlen, die Nutzung von Evidenzen in der Planung von Vorhaben in die GIZ-internen Handreichungen aufzunehmen.** Die Auswertung hat gezeigt, dass die Nutzung von Evidenzen in der Planung von Neu- und/oder Folgevorhaben in den betrachteten Dokumenten der GIZ (zum Beispiel Handreichungen) weder empfohlen noch vorgeschrieben ist. Der Begriff Evidenz wird hier breit gefasst und kann sowohl Evaluierungsergebnisse als auch anderweitige Erkenntnisse, zum Beispiel im Rahmen von akademischen Studien, beinhalten. Bisher beruhen Planungsprozesse, laut den Interviews, hauptsächlich auf der Nutzung der Erfahrungen derjenigen, die in die Planung involviert sind. Die Nutzung der Evidenzen sollte daher im Planungs- und Prüfprozess, wenn nicht verankert, zumindest als Empfehlung enthalten sein.

**Der GIZ wird empfohlen, einen unternehmensweiten Prozess zu etablieren, in dem die aufbereiteten und systematisierten Ergebnisse der Evaluierungen der GIZ (beispielsweise sektoral) im FMB und anderen relevanten Organisationseinheiten verankert werden.** Die Auswertung hat gezeigt, dass Evidenzen

bislang eher selten im Planungsprozess genutzt werden. Dabei wurde ebenfalls deutlich, dass die Planung von Neu- und/oder Folgevorhaben meist nur auf dem Erfahrungsschatz derjenigen, die in die Planung involviert sind, beruht. In diesem Rahmen ist in der Auswertung auch die fehlende Vergemeinschaftung der Evaluationserkenntnisse durch ZPE, QSA oder USE in der GIZ deutlich geworden, zum Beispiel von ZPE-Ergebnissen im gleichen Sektor oder Land wie das (geplante) Vorhaben. Der GIZ wird daher empfohlen, einen unternehmensweiten Prozess zu etablieren, der zum Ziel hat, die aufbereiteten und systematisierten Ergebnisse im FMB und anderen relevanten Organisationseinheiten zu verankern. Dies kann beispielsweise durch die Einführung eines ZPE-Dashboards geschehen, das es den in die Planung involvierten Akteur\*innen ermöglicht, in der Beratung / Planung von Vorhaben auf diese Evidenzen zurückzugreifen. Die Nutzung des Dashboards kann zudem gegebenenfalls durch Nennung in den Angeboten und/oder Prüfungen verankert werden.

## Literatur

BMZ (2019a): *Handbuch der bilateralen EZ: Handreichung Nr. HR060: Handreichung und kommentierte Gliederung für Modulvorschläge*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2019b): *Handbuch der bilateralen EZ: Handreichung Nr. HR013: Handreichung und kommentierte Gliederung für Kurzstellungnahmen für FZ/TZ-Module*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2019c): *Handbuch der bilateralen EZ: Handreichung Nr. HR069: Handreichung und kommentierte Gliederung für Modulvorschläge der Sektorvorhaben der TZ sowie IZR-Vorhaben*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2020a): *Handbuch der bilateralen EZ: Handreichung Nr. HR067: Verfahrensablauf für die Planung und Durchführung von Sektorvorhaben der TZ und Vorhaben für die Internationale Zusammenarbeit mit Regionen (IZR)*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2020b): *Handbuch der bilateralen EZ: Verfahrensinformation Nr. VI069: Fragenkatalog zur Gemeinsamen Verfahrensreform (GVR) - Frequently Asked Questions (FAQ)*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2020c): *Handbuch der bilateralen EZ: Verfahrensinformation Nr. VI083: Qualitätsstandards für Indikatoren*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

BMZ (2021): *Handbuch der bilateralen EZ: Handreichung Nr. HR013: Erstellung von Länder- und Regionalstrategien*, Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, unveröffentlichtes Dokument.

Bundesrechnungshof (2021): *Abschließende Mitteilung an das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung über die Prüfung Evaluierung von Maßnahmen der Entwicklungszusammenarbeit: Teil II – Maßnahmen der bilateralen Technischen Zusammenarbeit*, Potsdam: Bundesrechnungshof [online] <https://www.bundesrechnungshof.de/de/veroeffentlichungen/produkte/pruefungsmittelungen/2021/evaluierung-von-massnahmen-der-entwicklungszusammenarbeit-teil-ii-massnahmen-der-bilateralen-technischen-zusammenarbeit> [13.06.2022].

GIZ (2014): *Indikatoren: Eine Arbeitshilfe*, Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, unveröffentlichtes Dokument.

GIZ (2020): *Methodisches Vorgehen für die Prüfung: Arbeitshilfe*, Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, unveröffentlichtes Dokument.

GIZ (n.d.): *Evaluability Assessment: Final Checklist*, Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, unveröffentlichtes Dokument

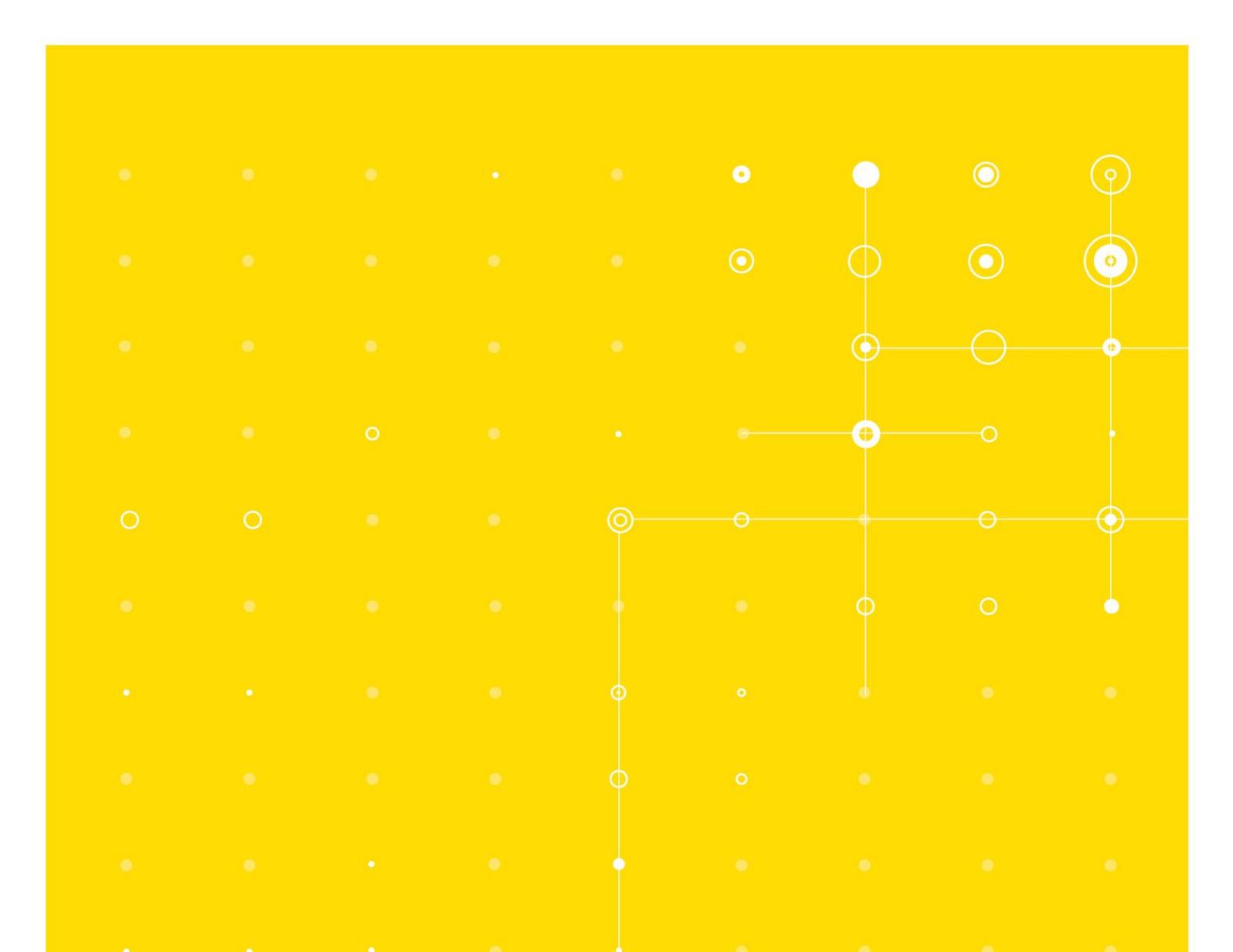
## **Anhang**

### **Anhang 1: Analyseraster**

Separater Anhang

### **Anhang 2: Erhebungsmaske**

Separater Anhang



## Fotonachweise und Quellen

### URL-Verweise:

In dieser Publikation befinden sich ggf. Verweise zu externen Internetseiten. Für die Inhalte der aufgeführten externen Seiten ist stets der jeweilige Anbieter verantwortlich. Die GIZ hat beim erstmaligen Verweis den fremden Inhalt daraufhin überprüft, ob durch ihn eine mögliche zivilrechtliche oder strafrechtliche Verantwortlichkeit ausgelöst wird. Eine permanente inhaltliche Kontrolle der Verweise auf externe Seiten ist jedoch ohne konkrete Anhaltspunkte einer Rechtsverletzung nicht zumutbar. Wenn die GIZ feststellt oder von anderen darauf hingewiesen wird, dass ein externes Angebot, auf das sie verwiesen hat, eine zivil- oder strafrechtliche Verantwortlichkeit auslöst, wird sie den Verweis auf dieses Angebot unverzüglich aufheben. Die GIZ distanziert sich ausdrücklich von derartigen Inhalten.

### Kartenmaterial:

Kartografischen Darstellungen dienen nur dem informativen Zweck und beinhalten keine völkerrechtliche Anerkennung von Grenzen und Gebieten. Die GIZ übernimmt keinerlei Gewähr für die Aktualität, Korrektheit oder Vollständigkeit des bereitgestellten Kartenmaterials. Jegliche Haftung für Schäden, die direkt oder indirekt aus der Benutzung entstehen, wird ausgeschlossen.



Deutsche Gesellschaft für  
Internationale Zusammenarbeit (GIZ) GmbH

Sitz der Gesellschaft  
Bonn und Eschborn

Friedrich-Ebert-Allee 32 + 36	Dag-Hammarskjöld-Weg 1-5
53113 Bonn, Deutschland	65760 Eschborn, Deutschland
T +49 228 44 60-0	T +49 61 96 79-0
F +49 228 44 60-17 66	F +49 61 96 79-11 15

E [info@giz.de](mailto:info@giz.de)  
I [www.giz.de](http://www.giz.de)