

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Schritt 3: Wirkungsmessung planen

Schon vor der Umsetzung des Projekts muss geplant werden, wie die Wirkung später gemessen und beurteilt werden soll. Die Ergebnisse dieses Schrittes werden schriftlich festgehalten.

So wird es gemacht

Aktivitäten Es muss geklärt werden, was wem wozu beantwortet werden soll. Daraus kann abgeleitet werden, womit die Resultate der Wirkungsmessung verglichen werden und wer die Evaluation später durchführt. Die angestrebten Leistungen und Wirkungen müssen in Indikatoren überführt werden und es braucht Zielwerte für diese Messgrössen. Für jeden Indikator muss geplant werden, mit welchen Methoden die nötigen Daten erhoben werden.

Fragen Mit der Beantwortung folgender Fragen kann der dritte Schritt der Wirkungsmessung angegangen werden:

- Was wollen wir in Erfahrung bringen?
- Womit werden die Ergebnisse verglichen?
- Wer soll die Wirkungsmessung durchführen und wer ist dafür verantwortlich?
- Mit welchen Indikatoren können wir unsere Outputs (Leistungen) und Outcomes (Wirkungen) messen?
- Aus welchen Quellen können die Daten bezogen werden?
- Wie werden die Daten erhoben und wer ist dafür zuständig?
- Wurde bei der Formulierung des Auftrags zur Wirkungsmessung an alles gedacht?

Resultate

- Verantwortlichkeiten stehen fest.
- Indikatoren, Datenquellen, Erhebungsmethoden, Häufigkeit und Zeitpunkt der Messung sowie Bezugsgrössen für Vergleiche sind klar.
- Konzept, Plan und Auftrag zur Wirkungsmessung sind formuliert.

Hilfsmittel

- Logical Framework Approach: Logical Framework Matrix
- Outcome Mapping: Outcome & Performance Monitoring, Evaluation Planning
- Theory of Change: Developing Indicators

Beispiele

- Brunnenbau: Ziele, Indikatoren, Messung
- Gesundheitskurs: Ziele, Indikatoren, Messung
- Medizinische Versorgung: Ziele, Indikatoren, Messung

WICHTIG

Innerhalb dieses Schrittes und zu den vorangegangenen Schritten (Projektziel definieren und Wirkungsmodell entwickeln) gibt es wichtige Feedbackschlaufen:

- Die Formulierung der Wirkungsziele mittels Indikatoren hilft, zu überprüfen, ob die geplanten Projektziele realistisch und erreichbar sind.
- Die Planung der Datenerhebung hilft, zu überprüfen, ob die Indikatoren rechtzeitig und mit vertretbarem Aufwand gemessen werden können.
- Beides zusammen zeigt, ob die Evaluationsfrage beantwortet werden kann.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme

Vergleiche

Die Planung einer Messung beinhaltet immer auch die Definition des Massstabes zur Bewertung des unter Gegenstandes. Die Definition des Vergleichsmassstabes bestimmt weitgehend, wie die Untersuchung ange werden sollte. Dabei müssen folgende Dimensionen beachtet werden:

- **Zeit**
Die Messung kann vor dem Projekt, am Ende oder nach Abschluss des Projekts durchgeführt werden. Soll der Verlauf des Projekts untersucht werden, sind zusätzliche Messungen während des Projekts nötig.
- **Bezugsgrösse**
Der Ist-Zustand kann mit den Zielen, der Ausgangslage oder mit einer Kontrollgruppe verglichen werden.

Basismodelle

Mit folgenden Basismodellen können die Zielerreichung, die Veränderungen bei der Zielgruppe oder der Ei eines Projekts gemessen werden:

Soll-Ist Vergleich	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Ziele		Definition	
Zielgruppe		Messung 1	
Vorher-Nachher Vergleich	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Zielgruppe	Messung 1	Messung 2	
Vergleich mit Kontrollgruppe	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Zielgruppe		Messung 1	

Kontrollgruppe

Messung 2

Vorher-Nachher
Vergleich mit
Kontrollgruppe

**vor dem Projekt
(Ex-ante)**

**am Ende des
Projekts**

**nach Abschluss
(Ex-post)**

Zielgruppe

Messung 1

Messung 3

Kontrollgruppe

Messung 2

Messung 4

WICHTIG

Eine rigorose Wirkungsmessung kann im Prinzip nur mit Vorher-Nachher Vergleich mit Kontrollgruppe werden. Festgestellte Veränderungen können so eindeutig dem Projekt zugeschrieben und externe I ausgeschlossen werden. Solche Vergleichsmodelle werden in der Praxis zunehmend von institutionen verlangt. Sie sind aber sehr aufwändig: Es braucht mindestens vier Messungen. Die Auswahl der Kc das Ziehen von Stichproben ist keineswegs trivial. Insbesondere kleinere Organisationen werden ka Ressourcen haben, rigorose Wirkungsmessungen durchzuführen und auch grössere Organisationer überlegen, wann und wo es für sie Sinn macht.

An dieser Stelle wird dafür plädiert, dass auch einfachere Modelle zweckmässig sind. Sie eignen sic für organisationsinternes Lernen. Damit Vorher-Nachher Vergleiche möglich sind, sollte jedoch eine Beginn des Projekts (Baselinestudie) eingeplant werden. Damit Soll-Ist Vergleiche gemacht werden klare Ziele formuliert sein.

Für die kausale Zuordnung (Attribution) der Wirkung zu einer Intervention braucht es rigorose Metho Wirkungsmessung. Zur Plausibilisierung, dass eine Intervention einen Beitrag zur Erreichung überge geleistet hat (Contribution), genügen auch einfachere Ansätze. In jedem Fall muss aber klar sein, wi Wirkungsmessung gemacht wurde und welche Schlüsse sich daraus ableiten lassen.

ACHTUNG

In der Praxis kommt es immer wieder vor, dass lediglich der Endzustand eines Projekts beschrieben Messung der Wirkung reicht eine Beschreibung der Zielgruppe ohne Bezug zu den Zielsetzungen, d Ausgangslage oder zu einer Kontrollgruppe jedoch nicht aus.

Erweiterte Modelle

Komplexere Modelle mit erhöhter Aussagekraft entstehen durch die Erweiterung und die Kombination von Basismodellen.

Vorher-Nachher und
Soll-Ist Vergleich

**vor dem Projekt
(Ex-ante)**

**am Ende des
Projekts**

**nach Abschluss
(Ex-post)**

Ziele

Definition

Zielgruppe	Messung 1	Messung 2	
Ex-post Vergleich	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Zielgruppe	Messung 1	Messung 2	Messung 3
Ex-post und Soll-Ist Vergleich	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Ziele		Definition	
Zielgruppe	Messung 1	Messung 2	Messung 3
Ex-post Vergleich mit Kontrollgruppe	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Zielgruppe	Messung 1	Messung 3	Messung 5
Kontrollgruppe	Messung 2	Messung 4	Messung 6
Soll-Ist Vergleich mit Kontrollgruppe	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Ziele		Definition	
Zielgruppe		Messung 1	
Kontrollgruppe		Messung 2	
Vorher-Nachher und Soll-Ist Vergleich mit Kontrollgruppe	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)
Ziele		Definition	
Zielgruppe	Messung 1	Messung 3	
Kontrollgruppe	Messung 2	Messung 4	
Ex-post- und Soll-Ist Vergleich mit Kontrollgruppe	vor dem Projekt (Ex-ante)	am Ende des Projekts	nach Abschluss (Ex-post)

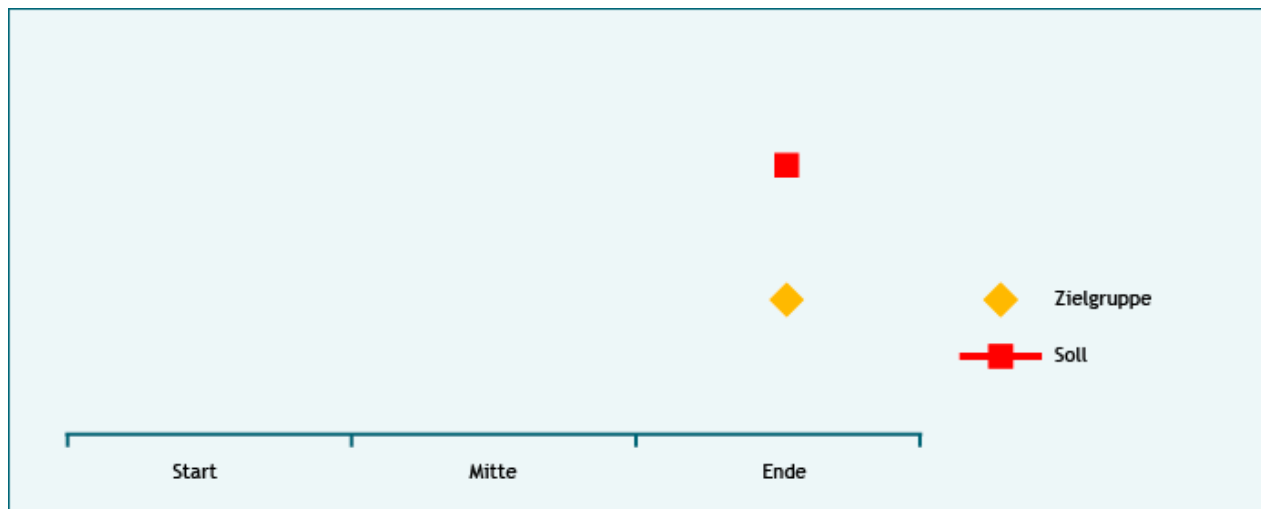
Ziele		Definition	Definition
Zielgruppe	Messung 1	Messung 3	Messung 5
Kontrollgruppe	Messung 2	Messung 4	Messung 6

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Soll-Ist Vergleich



Charakteristik

Es handelt sich um eine statische Betrachtung, bei der zu einem bestimmten Zeitpunkt die tatsächliche Situation der Zielgruppe mit der zu diesem Zeitpunkt geplanten Situation der Zielgruppe verglichen wird.

Beispiel

40% der jungen Mütter in der Region kennen am Ende des Prozesses den Zusammenhang von sauberem Trinkwasser und Gesundheit. Das sind deutlich weniger als geplant (Soll):

Vorteile	Nachteile
Eine einfache Art, die Zielerreichung festzuhalten.	Es gibt keinen Bezug zur Ausgangslage: Man weiss nicht, ob und wie sich die Situation der Zielgruppe seit dem Beginn der Intervention verändert hat.
Der Aufwand für die Datenerhebung ist gering.	Man weiss nicht, ob und wie sich die Situation der Zielgruppe verändert hätte, ohne das Projekt.
Das methodisches Know-how ist intern vorhanden oder erlernbar.	Es ist keine Aussage zur Wirkung möglich.
	Es ist keine Aussage zur Nachhaltigkeit möglich.

Eignung

Ist bedingt geeignet für das Lenken der Organisation und das Lernen innerhalb einer Organisation. Geeignet zu

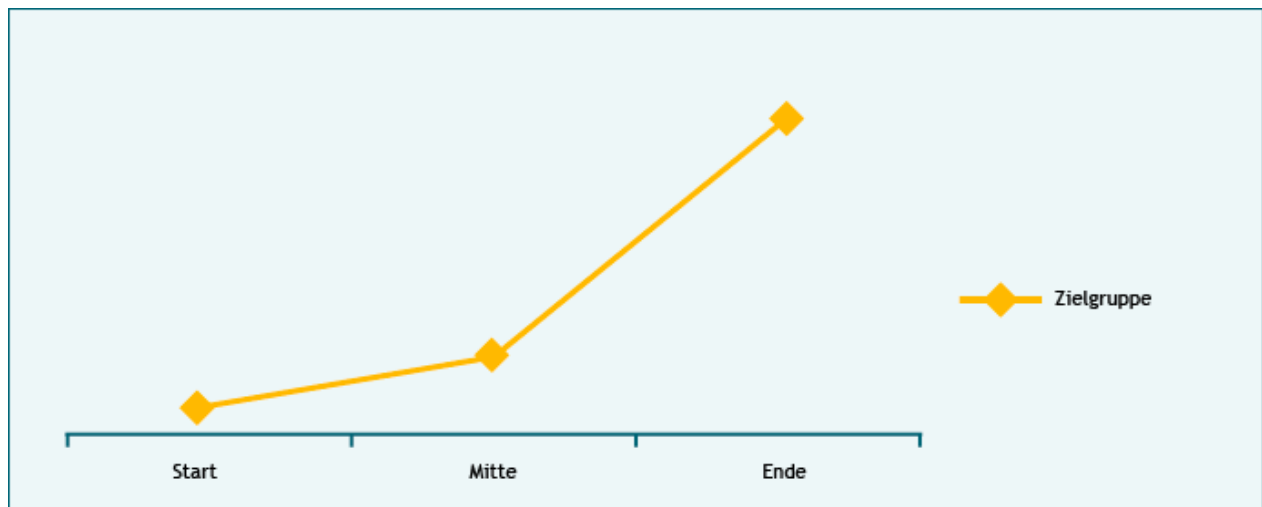
Legitimationszwecken bei Zielvereinbarungen. Kommt insbesondere dann zum Einsatz, wenn Daten zur Ausgangslage und geeignete Kontrollgruppen fehlen oder nur mit unverhältnismässig grossem Aufwand beschafft werden können. Zur Wirkungsmessung ist ein punktueller Soll-Ist Vergleich nicht geeignet.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Vorher-Nachher Vergleich



Charakteristik

Es wird die Entwicklung der Zielgruppe eines Projekts oder Programms beschrieben, z.B. am Anfang und am Ende des Projekts. Zusätzliche Messungen geben Auskunft über den Verlauf des Projekts.

Beispiel

70% der ländlichen Bevölkerung in der Region x haben innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser. Zu Beginn des Projekts hatten nur 10% der ländlichen Bevölkerung innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser.

Vorteile

Es kann plausibel gemacht werden, ob eine Intervention einen Beitrag zur angestrebten Wirkung geleistet hat oder nicht.

Aufwand für die Datenerhebung ist oft vertretbar.

Methodisches Know-how ist oft intern vorhanden oder erlernbar.

Nachteile

Man weiss nicht, ob und wie sich die Situation der Zielgruppe verändert hätte ohne das Projekt. Eine eindeutige Zuordnung der Wirkung zur Intervention ist nicht möglich

Es gibt keinen Bezug zur Zielsetzung. Man weiss nicht, ob und inwieweit die Organisation die gesetzten Ziele erreicht hat.

Aussagen zur Nachhaltigkeit des Projekts sind nicht möglich.

Eignung

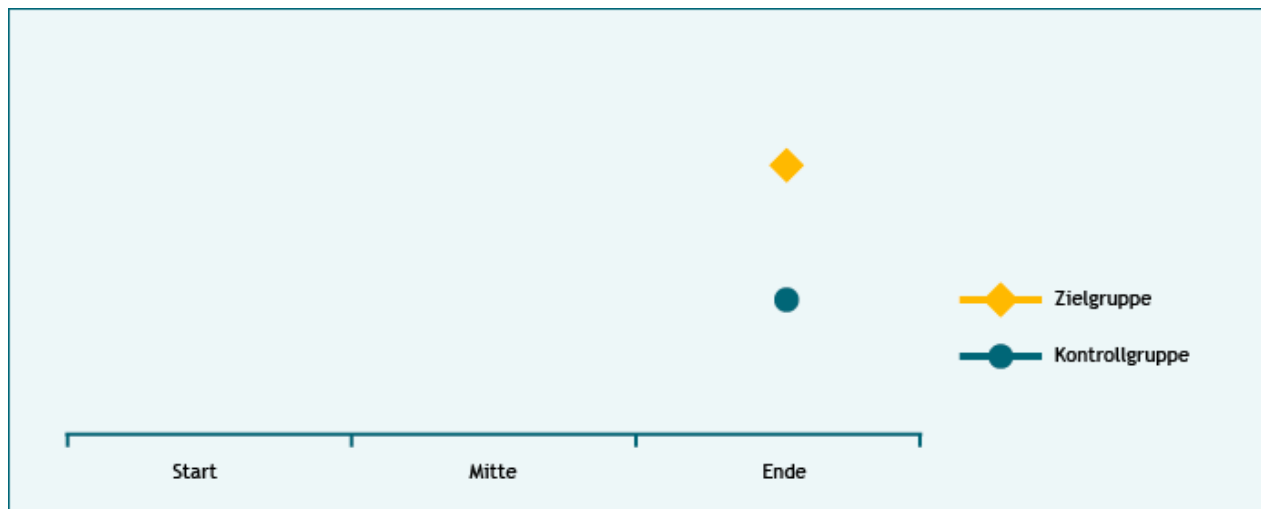
Geeignet, um aus den beobachteten Veränderungen zu lernen sowie zur Legitimation, wenn eine Plausibilisierung der Wirkung genügt. Kommt dann zum Einsatz, wenn Daten zur Ausgangslage vorhanden sind oder mit vertretbarem Aufwand rekonstruiert werden können, jedoch geeignete Kontrollgruppen fehlen oder nur mit unverhältnismässig grossem Aufwand konstruiert werden können oder wenn ein Vergleich mit Kontrollgruppen ethisch fragwürdig sind.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Vergleich mit Kontrollgruppe



Charakteristik

Es handelt sich um eine statische Betrachtung, bei der zu einem bestimmten Zeitpunkt die Situation der Zielgruppe mit der Situation der Kontrollgruppe verglichen wird.

Beispiel

90% der an Durchfall erkrankten Kinder in der Region x konnten erfolgreich medizinisch behandelt werden. In der Kontrollgruppe der Region y konnten nur 60% der an Durchfall erkrankten Kinder erfolgreich behandelt werden.

Vorteile

Es kann plausibel gemacht werden, ob eine Intervention einen Beitrag zur angestrebten Wirkung geleistet hat oder nicht.

Falls anfänglich identische Kontrollgruppen vorhanden sind, können Aussagen zur Wirkung des Projekts bei der Zielgruppe gemacht werden.

Der Aufwand für die Datenerhebung (2 Messungen) ist oft noch vertretbar.

Nachteile

Es gibt keinen Bezug zur Ausgangslage: Man weiss nicht, ob und wie sich die Situation der Ziel- und der Kontrollgruppe seit dem Beginn der Intervention verändert hat.

Identische Gruppen fehlen oft, methodisches Know-how ist erforderlich, wenn Kontrollgruppen konstruiert werden müssen. Die Wahl der Kontrollgruppe ist nicht trivial.

Es gibt keinen Bezug zu den gesetzten Zielen: Man weiss nicht, ob und inwieweit die Organisation ihre Ziel erreicht hat.

Es sind keine Aussagen zur Nachhaltigkeit möglich

Eignung

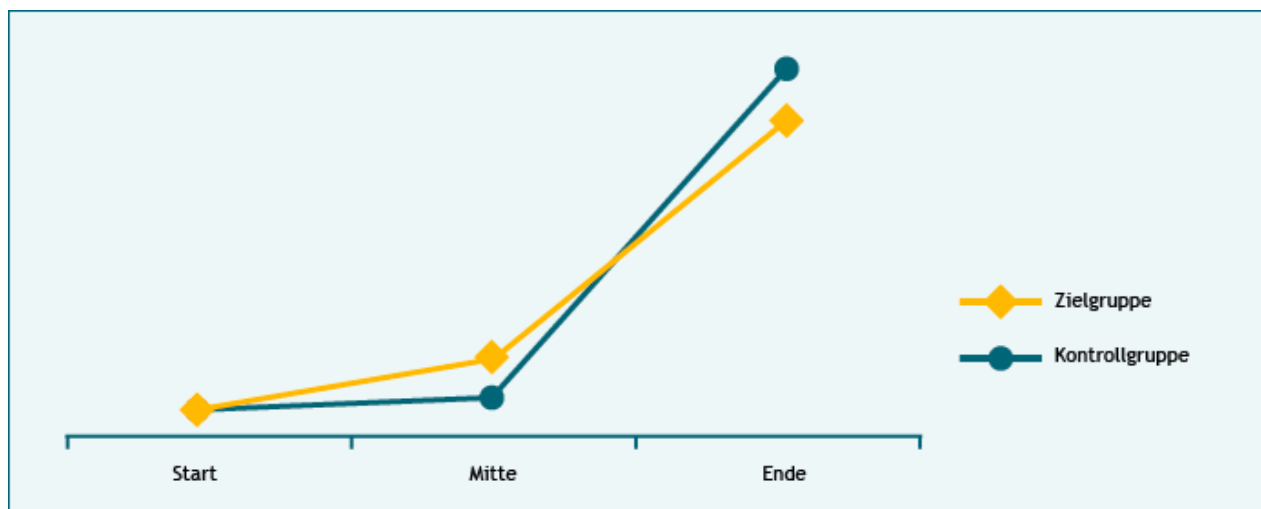
Durch seine erhöhte Aussagekraft ist ein Vergleich mit Kontrollgruppe für Legitimationszwecke geeignet, sofern geeignete Kontrollgruppen vorhanden sind oder konstruiert werden können und es keine ethischen Bedenken gibt. Er wird dann angewandt, wenn Daten zur Ausgangslage fehlen oder nur mit unverhältnismässig grossem Aufwand beschafft werden können.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Vorher-Nachher Vergleich mit Kontrollgruppe



Charakteristik

Die Entwicklung der Zielgruppe wird mit jener der Kontrollgruppe verglichen. Zusätzliche Messungen geben Auskunft über den Verlauf des Projekts.

Beispiel

90% der an Durchfall erkrankten Kinder konnten erfolgreich medizinisch behandelt werden. Das ist zwar bedeutend mehr als zu Beginn des Projekts. Die Situation der Kontrollgruppe hat sich ohne Intervention in der selben Zeit aber noch stärker verbessert.

Vorteile

Der Intervention kann eine bestimmte Wirkung zu- oder abgesprochen werden.

Der Ansatz ist methodisch fundiert.

Nachteile

Die Datenerhebung ist aufwändig und erfordert mindestens 4 Messungen.

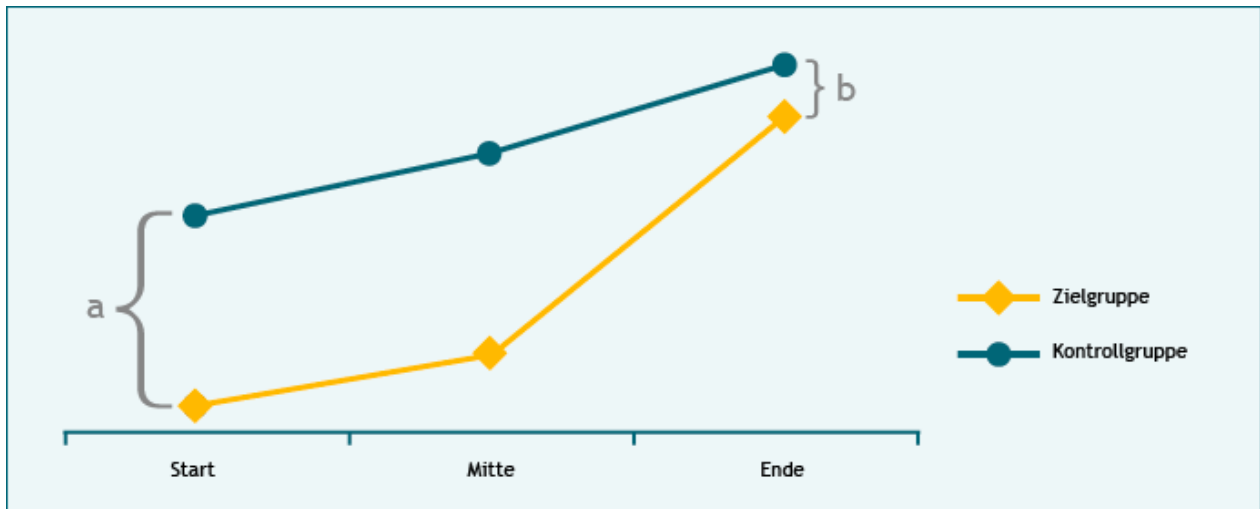
Das Vorgehen ist methodisch anspruchsvoll.

Die Zielerreichung wird nicht gemessen.

Es ist keine Aussage zur Nachhaltigkeit möglich.

Eignung

Eignet sich sehr gut zum Legitimieren, aber auch zum Lernen und Lenken der Organisation. Wird dann angewandt, wenn die Wirkung eindeutig nachgewiesen werden soll und es keine ethischen Bedenken zum Vergleich mit Kontrollgruppen gibt. Geeignete Kontrollgruppen müssen bestimmt oder konstruiert werden können.

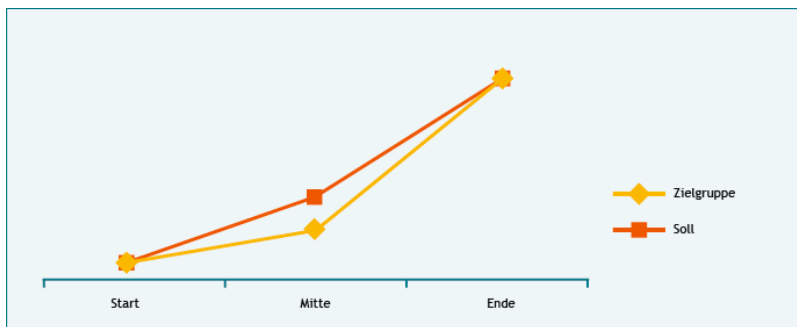


Wenn die Ausgangslage der Kontrollgruppe nicht mit jener der Zielgruppe identisch ist, muss die Differenz zwischen der Ziel- und der Kontrollgruppe am Anfang a) und am Ende b) des Projektes bestimmt werden. Die Aussage zur Wirkung hängt in diesem Fall davon ab, ob sich die Differenz vergrößert oder verringert hat. Dieses Verfahren wird difference-in-difference genannt.

Wirkungsmessung in der Entwicklungszusammenarbeit
Zewo-Leitfaden für Projekte und Programme



Vorher-Nachher und Soll-Ist Vergleich



Charakteristik

Die tatsächliche Entwicklung der Zielgruppe wird mit der geplanten Entwicklung der Zielgruppe verglichen. Zusätzliche Messungen geben Auskunft über den Verlauf des Projekts.

Beispiel

80% der jungen Mütter in der Region x kennen den Zusammenhang von sauberem Trinkwasser und Gesundheit. Das sind dreimal mehr als zu Beginn des Projekts und so viele wie angestrebt.

Vorteile

Es wird gezeigt, inwieweit die gesetzten Ziele erreicht werden konnten.

Es kann plausibel gemacht werden, ob eine Intervention einen Beitrag zur angestrebten Wirkung geleistet hat oder nicht.

Der Aufwand für die Datenerhebung (2 Messungen) ist oft vertretbar.

Das methodische Know-how ist oft vorhanden oder erlernbar.

Nachteile

Man weiss nicht, ob und wie sich die Situation der Zielgruppe verändert hätte ohne das Projekt. Eine eindeutige Zuordnung der Wirkung zur Intervention ist nicht möglich.

Es ist keine Aussage zur Nachhaltigkeit möglich.

Eignung

Ist zum Lernen und insbesondere zum Lenken der Organisation geeignet kann aber bei Zielvereinbarungen der wenn eine Plausibilisierung der Wirkung genügt, auch zu Legitimation dienen. Kommt zum Einsatz, wenn klare Ziele formuliert wurden und Daten zur Ausgangslage vorhanden sind, jedoch geeignete Kontrollgruppen fehlen oder wenn Vergleiche mit Kontrollgruppen ethisch fragwürdig sind.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Ex-post Vergleich



Charakteristik

Es wird die Entwicklung der Zielgruppe über die Beendigung des Projekts hinaus betrachtet.

Beispiel

Drei Jahre nach Beendigung des Projekts haben 70% der ländlichen Bevölkerung in der Region x innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser. Das sind mehr als vor dem Projekt, aber weniger als bei Projektende.

Vorteile

Der Aufwand für die Datenerhebung ist oft vertretbar.

Das methodische Know-how ist oft intern vorhanden oder erlernbar.

Es kann plausibel gemacht werden, ob eine Intervention einen Beitrag zur angestrebten Wirkung geleistet hat oder nicht.

Es kann gezeigt werden, ob eine Intervention nachhaltig war oder nicht.

Nachteile

Externe Einflussfaktoren können nicht ausgeschlossen werden, da der Vergleich zur Kontrollgruppe fehlt.

Es wird nicht ersichtlich, ob die gesetzten Ziele erreicht worden sind.

Die Erkenntnisse liegen erst Jahre nach Abschluss der Intervention vor.

Eignung

Kommt dann zum Einsatz, wenn die Nachhaltigkeit einer Intervention untersucht werden soll. Durch die zeitliche Verzögerung kommt die Information zur kurzfristigen Legitimation und für das unmittelbare Lenken und Lernen zu spät. Sie kann jedoch für die langfristige Entwicklung der Organisation und die strategische Ausrichtung genutzt werden.

Eignung

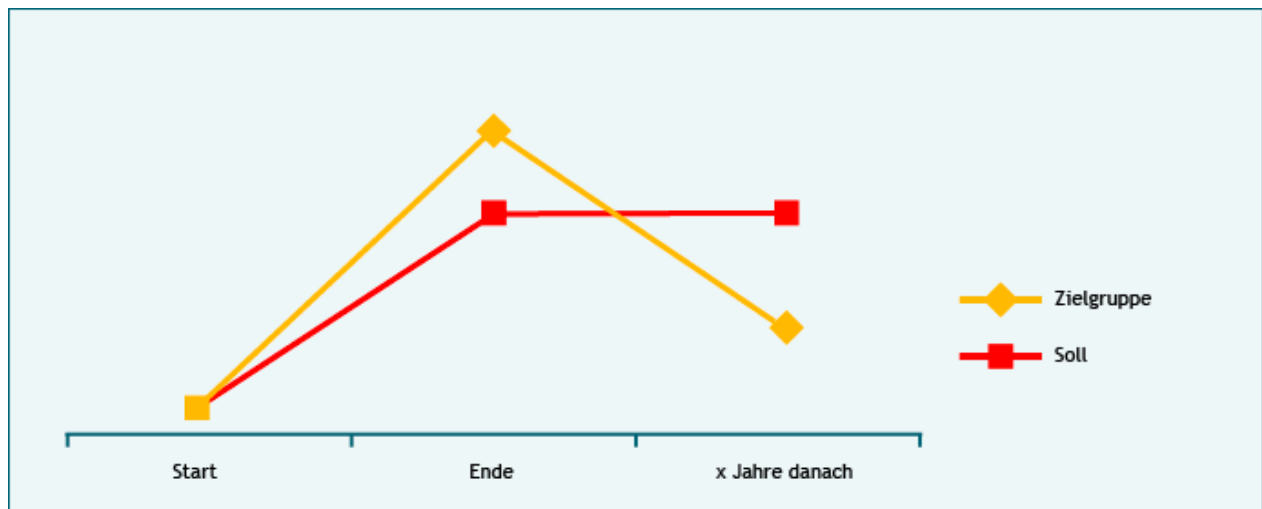
Eignet sich zur Legitimation bei Zielvereinbarungen oder wenn eine Plausibilisierung der Wirkung genügt. Kommt zum Einsatz, wenn die Nachhaltigkeit einer Intervention untersucht werden soll und Vergleiche mit Kontrollgruppen zu aufwändig oder ethisch fragwürdig sind. Durch die zeitliche Verzögerung kommt die Information zur unmittelbaren Legitimation und für das kurzfristige Lenken und Lernen zu spät. Sie kann jedoch für die langfristige Entwicklung der Organisation und die strategische Ausrichtung genutzt werden.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Ex-post und Soll-Ist Vergleich



Charakteristik

Die tatsächliche Entwicklung der Zielgruppe wird mit der geplanten Entwicklung über das Projektende hinaus verglichen.

Beispiel

Drei Jahre nach Beendigung der Intervention kennen 30% der jungen Mütter in der Region x den Zusammenhang von sauberem Trinkwasser und Gesundheit. Das sind mehr als vor dem Projekt, aber weniger als am Ende des Projekts und weniger als angestrebt.

Vorteile

Es wird gezeigt, inwieweit die gesetzten Ziele erreicht werden konnten.

Es kann plausibel gemacht werden, ob eine Intervention einen Beitrag zur angestrebten Wirkung geleistet hat oder nicht.

Es kann gezeigt werden, ob eine Intervention nachhaltig war oder nicht.

Das methodische Know-how ist vorhanden oder erlernbar.

Nachteile

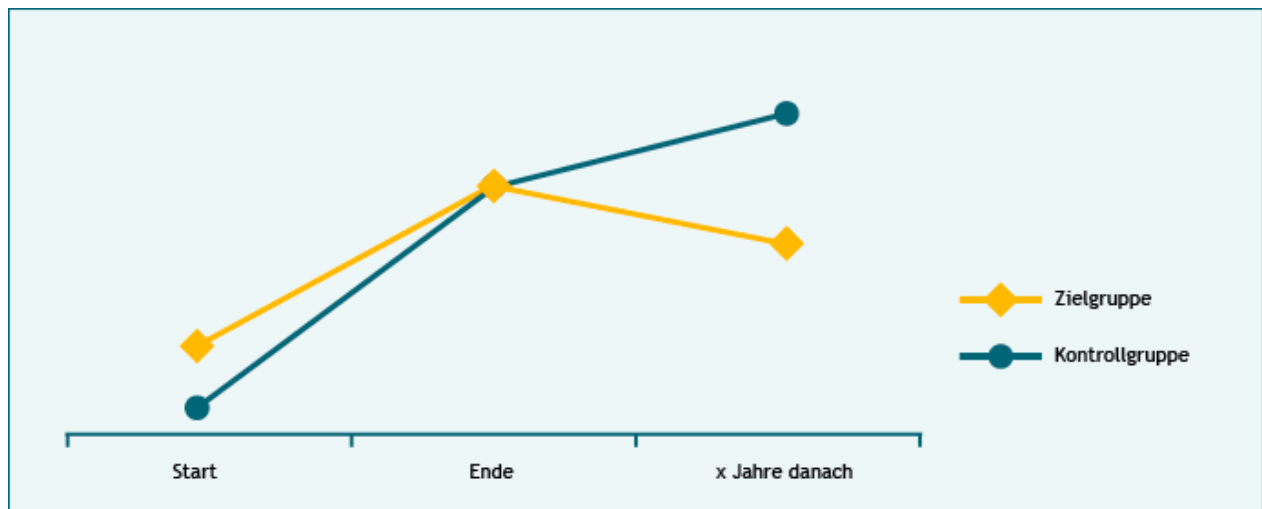
Man weiss nicht, ob und wie sich die Situation der Zielgruppe verändert hätte ohne das Projekt. Eine eindeutige Zuordnung der Wirkung zur Intervention ist nicht möglich.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Ex-post Vergleich mit Kontrollgruppe



Charakteristik

Die Entwicklung der Zielgruppe wird mit jener der Kontrollgruppe über das Projektende hinaus verglichen.

Beispiel

30% der jungen Mütter in der Region x kennen den Zusammenhang von sauberem Trinkwasser und Gesundheit. Das sind mehr als vor dem Projekt aber weniger als am Ende des Projekts. Der Wissensstand in der Kontrollgruppe hat sich in der gleichen Zeit von einer schwächeren Ausgangslage auf das Niveau der Zielgruppe entwickelt und verbessert sich weiter.

Vorteile

Es kann wissenschaftlich fundiert nachgewiesen werden, ob eine Intervention nachhaltig wirksam war oder nicht.

Nachteile

Die Datenerhebung ist sehr aufwändig und erfordert mindestens 6 Messungen.

Das Vorgehen ist methodisch anspruchsvoll.

Die Zielerreichung wird nicht gemessen.

Eignung

Geeignet, wenn wissenschaftlich fundiert nachgewiesen werden soll, ob mit der Intervention eine nachhaltige Wirkung erzielt worden ist. Kommt zum Einsatz, wenn es keine Bedenken gegen Vergleiche mit Kontrollgruppen gibt und geeignete Kontrollgruppen bestimmt oder konstruiert werden können. Die Daten zur Ausgangslage müssen allenfalls rekonstruiert werden. Wenn die Ausgangslage der Kontrollgruppe nicht mit jener der Zielgruppe identisch ist, muss die Differenz zwischen der Ziel- und der Kontrollgruppe verfolgt werden. Die Aussage zur Wirkung hängt in

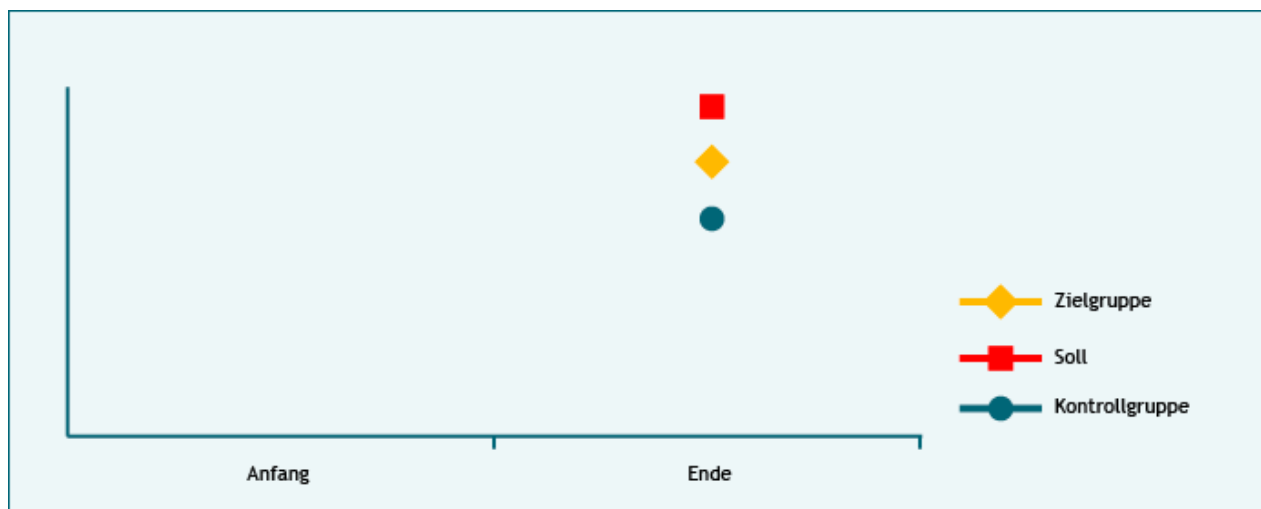
diesem Fall davon ab, ob sich die Differenz vergrößert oder verringert. Durch die zeitliche Verzögerung kommt die Information zur kurzfristigen Legitimation und für das unmittelbare Lenken und Lernen zu spät. Sie kann jedoch für die langfristige Entwicklung der Organisation und die strategische Ausrichtung genutzt werden.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Soll-Ist Vergleich mit Kontrollgruppe



Charakteristik

Es handelt sich um eine punktuelle Betrachtung, bei der zu einem bestimmten Zeitpunkt die Situation der Zielgruppe mit den Zielen und mit der Situation der Kontrollgruppe verglichen wird.

Beispiel

70% der jungen Mütter in der Region x kennen den Zusammenhang von sauberem Trinkwasser und Gesundheit. Das sind mehr als in der Kontrollgruppe, aber weniger als angestrebt.

Vorteile

Der Aufwand für die Datenerhebung mit 2 Messungen ist oft noch vertretbar.

Es ist eine einfache Art, die Zielerreichung festzuhalten.

Falls identische Kontrollgruppen vorhanden sind, können Aussagen zur Wirkung des Projekts bei der Zielgruppe gemacht werden.

Nachteile

Es gibt keinen Bezug zur Ausgangslage. Man weiss nicht, ob und wie sich die Situation der Ziel- und der Kontrollgruppe seit dem Beginn der Intervention verändert hat und ob sie vom selben Ausgangspunkt gestartet sind.

Identische Kontrollgruppen fehlen oft.

Es sind keine Aussagen zur Nachhaltigkeit möglich.

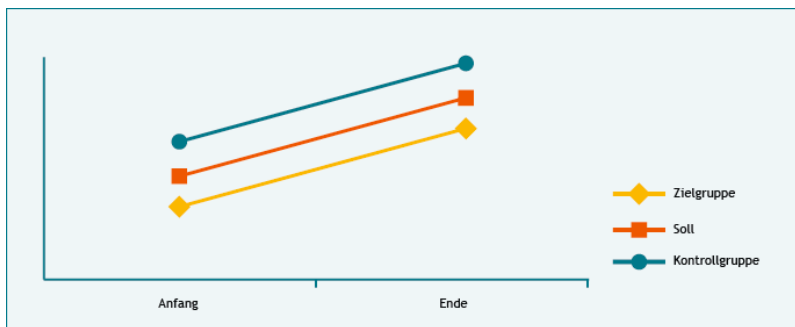
Eignung

Ist nur bedingt geeignet für das Lenken und Lernen innerhalb einer Organisation. Kann jedoch zur Legitimation bei Zielvereinbarungen geeignet sein. Wird in der Praxis dann angewandt, wenn Daten zur Ausgangslage fehlen oder nur mit unverhältnismässig grossem Aufwand beschafft werden können, jedoch geeignete Kontrollgruppen vorhanden sind.

Wirkungsmessung in der Entwicklungszusammenarbeit
Zewo-Leitfaden für Projekte und Programme



Vorher-Nachher und Soll-Ist Vergleich mit Kontrollgruppe



Charakteristik

Die Entwicklung der Zielgruppe wird mit jener der Kontrollgruppe verglichen und zur Zielsetzung in Bezug gesetzt wird.

Beispiel

50% der ländlichen Bevölkerung in der Region x hat innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser. Zu Beginn des Projekts waren es nur 30%. Die Verbesserung um 20%-Punkte ist gleich gut wie in der Kontrollgruppe in der Region y, bei der sich der Zugang zu Trinkwasser im selben Zeitraum von 50 auf 70% der Bevölkerung verbessert hat. Das Ziel von 60% konnte nicht erreicht werden.

Vorteile

Es wird gezeigt, was sich bei der Zielgruppe durch die Intervention verändert hat.

Es wird festgestellt, ob die Ziele erreicht worden sind.

Der Ansatz ist methodisch fundiert.

Die Zuordnung der Wirkung zur Intervention ist möglich.

Nachteile

Die Datenerhebung ist aufwändig und erfordert mindestens 4 Messungen.

Das Vorgehen ist methodisch anspruchsvoll.

Es ist keine Aussage zur Nachhaltigkeit möglich.

Eignung

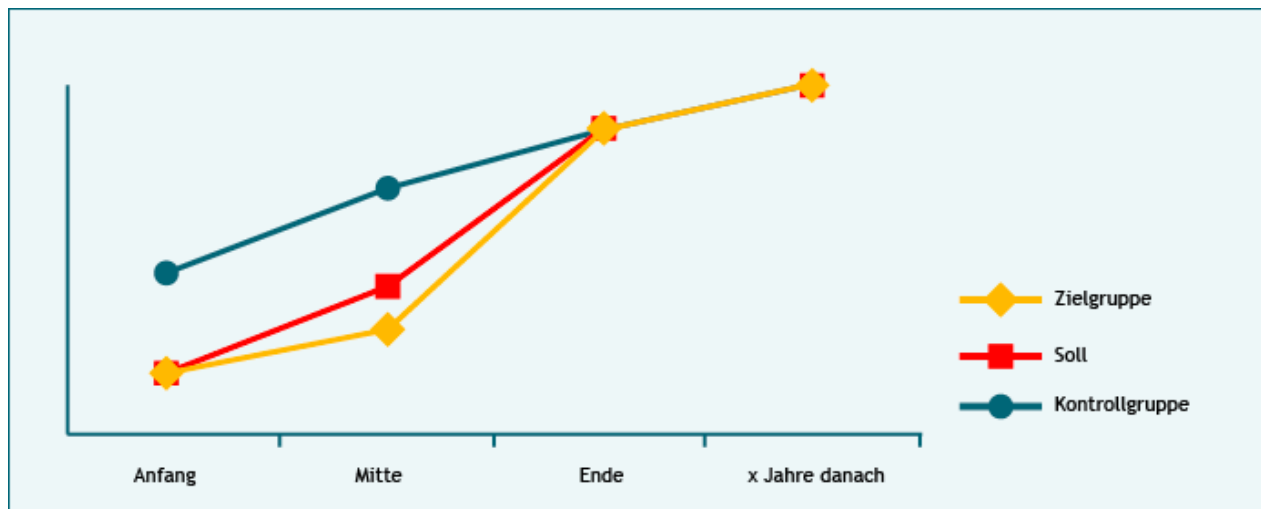
Eignet sich zum Lernen, Lenken und Legitimieren. Wird bei Zielvereinbarungen angewandt und insbesondere dann, eingesetzt, wenn die Wirkung eindeutig nachgewiesen werden soll und es keine ethischen Bedenken zum Vergleich mit Kontrollgruppen gibt. Geeignete Kontrollgruppen müssen bestimmt oder konstruiert werden können. Wenn die Ausgangslage der Kontrollgruppe nicht mit jener der Zielgruppe identisch ist, muss die Differenz zwischen der Ziel- und der Kontrollgruppe am Anfang und am Ende des Projektes bestimmt werden. Die Aussage zur Wirkung hängt in diesem Fall davon ab, ob sich die Differenz vergrößert oder verringert hat. Dieses Verfahren wird difference-in-difference genannt.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Ex-post und Soll-Ist Vergleich mit Kontrollgruppe



Charakteristik

Die Entwicklung der Zielgruppe wird mit jener der Kontrollgruppe über das Projektende hinaus verglichen und mit den Zielsetzungen in Bezug gebracht.

Beispiel

55% der an Durchfall erkrankten Kinder konnten erfolgreich medizinisch behandelt werden. Das sind 25%-Punkte mehr als vor dem Projekt. In der gleichen Zeitspanne hat sich die Situation der Kontrollgruppe nur um 15%-Punkte verbessert. Nach Beendigung des Projekts verläuft die Entwicklung der Ziel- und der Kontrollgruppe gleich. Das Projekt hat die Soll-Ziele erreicht.

Vorteile

Es kann wissenschaftlich fundiert nachgewiesen werden, ob eine Intervention nachhaltig wirksam war und ob die Ziele langfristig erreicht wurden.

Nachteile

Die Datenerhebung ist sehr aufwändig (mind. 6 Messungen).

Das Vorgehen ist methodisch anspruchsvoll

Eignung

Eignet sich für die Überprüfung der langfristigen Entwicklungspolitik der Organisation und der strategischen Ausrichtung der Organisation. Geeignet für die Legitimation der nachhaltigen Wirkung gegenüber Dritten. Wegen der zeitlichen Verzögerung ist der Ansatz weniger geeignet für das unmittelbare Lenken und Lernen. Wird dann, eingesetzt, wenn eine nachhaltige Wirkung eindeutig nachgewiesen werden soll und es keine Bedenken zum

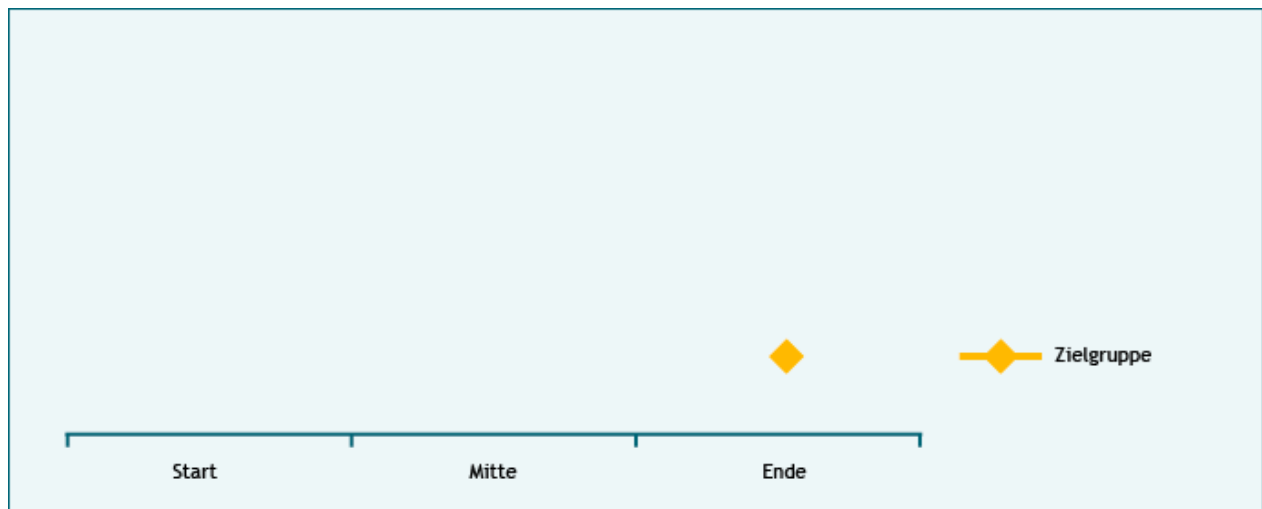
Vergleich mit Kontrollgruppen gibt. Geeignete Kontrollgruppen müssen bestimmt oder konstruiert werden können. Wenn die Ausgangslage der Kontrollgruppe nicht mit jener der Zielgruppe identisch ist, muss die Differenz zwischen der Ziel- und der Kontrollgruppe verfolgt werden. Die Aussage zur Wirkung hängt in diesem Fall davon ab, ob sich die Differenz vergrößert oder verringert.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beschreibung



Charakteristik

Es handelt sich um eine punktuelle Betrachtung, bei der zu einem bestimmten Zeitpunkt die Situation der Zielgruppe beschrieben wird.

Beispiele

70% der ländlichen Bevölkerung in der Region x haben innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser.

Vorteile	Nachteile
Der Aufwand für die Datenerhebung ist gering.	Es gibt keinen Bezug zu den gesetzten Zielen: Man weiss nicht, ob und inwieweit die Organisation ihre Ziele erreicht hat.
Das methodische Know-how ist vorhanden.	Es gibt keinen Bezug zur Ausgangslage: Man weiss nicht, ob und wie sich die Situation der Zielgruppe seit dem Beginn der Intervention verändert hat.
	Man weiss nicht, ob und wie sich die Situation der Zielgruppe verändert hätte, ohne das Projekt.
	Es ist keine Aussage zur Wirkung möglich.
	Es ist keine Aussage zur Nachhaltigkeit möglich.

Eignung

Durch ihre geringe Aussagekraft, ist die punktuelle Beschreibung wenig geeignet für das Lernen oder Lenken einer Organisation. In der Praxis wird insbesondere dann auf die punktuelle Beschreibung zurückgegriffen, wenn keine Ziele formuliert wurden, Daten zur Ausgangslage und Kontrollgruppen fehlen oder diese nur mit unverhältnismässig grossem Aufwand konstruiert werden können. Zur Wirkungsmessung ist die reine Beschreibung eines Zustandes jedoch nicht geeignet.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Planung und Eignung

Der Zeitpunkt der Messungen muss geplant werden. Die Häufigkeit der Messungen erhöht die Genauigkeit der Resultate, aber auch den Aufwand. Die folgende Tabelle zeigt, wann es welche Messungen braucht, wofür die sie geeignet sind und wie weit sie in der Praxis verbreitet sind.

Planung von Zeitpunkt und Häufigkeit der Messung

Zeitpunkt	Vor dem Projekt (Ex-ante)	Während dem Projekt	Am Ende des Projekts	Nach Abschluss des Projekts (Ex-post)
Notwendigkeit	Nötig, wenn erforderliche Daten zur Ausgangslage fehlen	Nötig, wenn die Verläufe betrachtet werden sollen	Nötig, wenn der Erfolg beurteilt werden soll	Nötig, wenn die nachhaltige Wirkung der Intervention beurteilt werden soll
Eignung	Geeignet, um über die Durchführung zu entscheiden	Geeignet, zur Überwachung und Steuerung der Umsetzung	Geeignet zur Steuerung und Rechenschaftsablage	Geeignet zur Überprüfung der Strategie und Politik
Verbreitung	Wird eher selten und wenn, dann vor grossen Programmen angewandt	Wird häufig nur auf Output-Ebene (Monitoring) angewandt	Wird auch auf Outcome-Ebene angewandt	Wird eher selten und wenn, dann auf Impact-Ebene angewandt

Die reine Beschreibung eines Zustandes sagt ebenso wenig über die Wirkung aus wie ein herkömmlicher Soll-Ist Vergleich. Damit Aussagen über die Wirkung möglich werden, muss der erreichte Zustand der Zielgruppe zu deren Ausgangslage oder zu einer Kontrollgruppe in Bezug gesetzt werden. Die Kombination von beidem wird gelegentlich auch als «gold standard» der Wirkungsmessung bezeichnet. Die nachfolgende Tabelle gibt einen Überblick, welche Vergleiche möglich sind und wozu sie sich eignen.

Eignung der Vergleichsarten

Aussagekraft	Ziel- erreichung	Wirkung	Plausibilisierung (Contribution)	Eindeutige Zuordnung (Attribution)	Nachhaltigkeit	Minimale Anzahl Messungen	Dauer
Basis Vergleichsmodelle							
Soll-Ist Vergleich	o.k.					1	punktuell
Vorher- Nachher		o.k.	o.k.			2	3 bis 8 Jahre

Vergleich							
Vergleich mit Kontrollgruppe		o.k.	o.k.			2	punktuell
Vorher-Nachher Vergleich mit Kontrollgruppe		o.k.	o.k.	o.k.		4	3 bis 8 Jahre
Komplexere Modelle und Kombinationen							
Vorher-Nachher und Soll-Ist Vergleich	o.k.	o.k.	o.k.			2	3 bis 8 Jahre
Ex-post Vergleich		o.k.	o.k.		o.k.	3	6 bis 12 Jahre
Ex-post und Soll-Ist Vergleich	o.k.	o.k.	o.k.		o.k.	3	6 bis 12 Jahre
Ex-post Vergleich mit Kontrollgruppe		o.k.	o.k.	o.k.	o.k.	6	6 bis 12 Jahre
Soll-Ist Vergleich mit Kontrollgruppe	o.k.	o.k.	o.k.			2	punktuell
Vorher-Nachher und Soll-Ist Vergleich mit Kontrollgruppe	o.k.	o.k.	o.k.	o.k.		4	3 bis 8 Jahre
Ex-post und Soll-Ist Vergleich mit Kontrollgruppe	o.k.	o.k.	o.k.	o.k.	o.k.	6	6 bis 12 Jahre
Kein Vergleich							
Beschreibung						1	punktuell

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Selbstevaluation

Die Selbstevaluation ist ein Verfahren, bei dem die praxisgestaltenden Fachleute identisch sind mit den Evaluatorinnen und Evaluatoren. Dies bedeutet, dass die Akteure ihre eigene Tätigkeit überprüfen: Sie sind also gleichzeitig praxisverantwortlich und evaluationsverantwortlich.¹ Inhaltlich unterscheiden sich die Fragestellungen der Selbstevaluation nicht von denjenigen einer externen Evaluation. Beurteilt werden insbesondere die Relevanz, die Wirksamkeit und die Wirtschaftlichkeit eines Projekts. Bei einer unterstützten Selbstevaluation berät und unterstützt ein anerkanntes Institut oder eine anerkannte Fachperson die Projektdurchführenden in der Planung, Durchführung und Berichterstattung einer Selbstevaluation.

Vorteile	Nachteile
Motivation der Evaluierten ist grösser	Weniger Distanz
Insiderwissen kann eingesetzt werden	Grundsatzfragen werden weniger gestellt
Evaluatoren sind mit dem Thema vertraut	Evaluationswissen fehlt
Ownership für Ergebnisse führt zu rascher Umsetzung	Weniger Legitimationskraft gegen aussen
Akzeptanz der Ergebnisse in der Organisation grösser	
Organisation ist weniger aufwändig	

Eignung

Die Selbstevaluation eignet sich besonders für Evaluationen, die gemacht werden, um daraus zu lernen. Sie eignet sich gut dafür Prozesse zu analysieren und schrittweise Verbesserungen zu erzielen. Selbstevaluationen können auch bei knappen finanziellen Ressourcen durchgeführt werden.

WICHTIG

Auch bei Selbstevaluationen ist darauf zu achten, dass dafür genügend Zeit und Ressourcen eingesetzt werden, dass das nötige Know-how vorhanden ist oder bereit gestellt wird und Verbindlichkeiten klar festgelegt sind. Selbstevaluationen scheitern sonst gern an der Methodenkomplexität und mangelnden Ressourcen.

¹ Quelle: Hildegard Müller-Kohlenberg, Wolfgang Beywl (2003): *Standards der Selbstevaluation, Begründung und aktueller Diskussionsstand*, Zeitschrift für Evaluation 1/2003, Köln.



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Durchführung

Die Durchführung der Wirkungsmessung kann in verschiedener Weise erfolgen. Es ist zu planen, durch wen die Daten erhoben und ausgewertet werden. Die Wirkungsmessung kann grundsätzlich von den Projektverantwortlichen selber (Selbstevaluation) oder von unabhängigen Dritten (Fremdevaluation) durchgeführt werden. Dazwischen gibt es verschiedene Mischformen (Hybridevaluation).

Dabei ist Folgendes zu beachten:

- **Unabhängigkeit**
Externe Experten oder auch Projektverantwortliche müssen, je nach Verwendungszweck der Wirkungsmessung (z.B. fürs Legitimieren oder Lenken), die nötige Unabhängigkeit haben, um von den Nutzerinnen und Nutzern der Ergebnisse als genügend unparteiisch und unvoreingenommen wahrgenommen zu werden.
- **Glaubwürdigkeit**
Das Team muss die nötige fachliche und methodische Kompetenz haben, um die Evaluation fachgerecht durchführen zu können.
- **Akzeptanz**
Das Team muss das nötige Einfühlungsvermögen und die nötige Erfahrung haben, um von den Beteiligten und den Nutzerinnen und Nutzern der Ergebnisse akzeptiert zu werden.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Hybridevaluation

Die Kombination von Selbst- und Fremdevaluation kann als Hybridevaluation bezeichnet werden. Dabei werden die Vorteile von Selbst- und Fremdevaluation kombiniert. Im Zentrum dieses Ansatzes steht ein systematischer, durch ein unabhängiges Evaluationsteam unterstützter Prozess einer Selbstevaluation. Das Evaluationsteam unterstützt das Programm bei der Selbstevaluation und der kritischen Beurteilung des Programms. Daten und Informationen, welche der Beurteilung zugrunde liegen, werden dabei vom Evaluationsteam verifiziert, zusätzliche Fragen werden aufgeworfen und alternative Hypothesen der Interpretation der Daten zur Diskussion gestellt. Dieses Vorgehen stellt sicher, dass die Evaluation eng auf die Bedürfnisse des Programteams abgestimmt und der distanzierte Blick dennoch gewährleistet ist.

Vorteile	Nachteile
Informationen können rechtzeitig ins Projekt einfließen	Rollen sind nicht immer klar
Lernorientierte Atmosphäre	Braucht Zeit, bis alle ihre Rolle kennen

Eignung

Hybridevaluationen eignen sich sowohl fürs Lernen als auch fürs Legitimieren. Bei knappen eigenen personellen Ressourcen oder wenn wenig internes Evaluations-Know-how vorhanden ist, aber trotzdem nicht auf die Vorteile einer Selbstevaluation verzichtet werden soll, ist die Hybridevaluation das geeignete Mittel.

WICHTIG

Die Zielsetzung und die Fragestellung der Evaluation müssen klar festgelegt werden. Die Auswahl geeigneter externer Evaluatoren ist entscheidend für das Gelingen der Evaluation. Der Verteilung der Rollen und der Festlegung der Verantwortlichkeiten sollte besondere Aufmerksamkeit geschenkt werden.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Fremdevaluation

Die Verantwortung für die Durchführung liegt bei Personen, welche nicht in die Umsetzung einer Massnahme einbezogen sind. Expertinnen und Experten verschaffen sich Informationen über den Evaluationsgegenstand und geben den Verantwortlichen Rückmeldungen darüber, wie sie diesen einschätzen.

Vorteile	Nachteile
Keine «Betriebsblindheit»	Geringe Einbindung in Arbeitsfeld
Unbefangenheit	Kontextbezug und Sachkompetenz fehlen
Methodenkompetenz	Weniger Legitimationskraft gegen innen
Akzeptanz gegen aussen ist höher	

Eignung

Die Fremdevaluation eignet sich besonders für Evaluationen, die gemacht werden, um ein Projekt zu legitimieren. Fremdevaluationen kommen zum Einsatz, wenn die eigenen personellen Ressourcen knapp sind oder wenn wenig internes Evaluations-Know-how vorhanden ist.

WICHTIG

Die Zielsetzung und die Fragestellung der Evaluation müssen klar festgelegt werden. Die Auswahl geeigneter externer Evaluatoren ist entscheidend für das Gelingen der Evaluation. Evaluatoren müssen natürlich die nötigen Sachkenntnisse mitbringen, aber auch über die nötige Sensibilität im Umgang mit verschiedenen Anspruchsgruppen verfügen und von diesen akzeptiert werden. Schliesslich sind realistische Zeitpläne und Budgets bereitzustellen.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Indikatoren

Damit Wirkungen gemessen werden können und das Erreichen von Wirkungszielen überprüft werden kann, müssen diese in objektiv messbaren Grössen konkretisiert werden. Eine solche objektiv messbare Grösse wird Indikator genannt. Ein Indikator gibt Antwort auf die Frage: «Wie werden wir wissen, ob das, was wir geplant haben, auch tatsächlich eingetroffen ist und dass wir unser Ziel erreicht haben?» Ein Indikator für ein Wirkungsziel gibt also an, wie wir die Wirkung messen und wann wir das Wirkungsziel als erreicht betrachten:

Ziel	Indikator
Verbesserter Zugang zu Trinkwasser	Nächster Brunnen ist innert 15 Gehminuten erreichbar
Kindersterblichkeit sinkt	Kindersterblichkeit sinkt von x% auf y%

WICHTIG

Ein guter Indikator sollte **SMART** sein:

- Specific: Der Indikator muss eindeutig und klar sein.
- Measurable: Der Indikator muss messbar sein, auch mit angemessenen Kosten.
- Achievable: Der vom Indikator vorgegebene Sollwert muss erreichbar sein.
- Relevant: Die Information die der Indikator liefert soll relevant für Projektmanager sein.
- Time-bound: Der Indikator muss angeben, bis wann das Ziel erreicht werden soll.

Quelle: *European Commission, PCM Guidelines*

Gute Indikatoren zu finden und zu wählen ist entscheidend dafür, dass eine Wirkungsmessung nützliche Informationen liefert, aber keineswegs einfach. Eine hohe Bedeutung kommt auch hier der partizipativen Erarbeitung zu: Ein guter Indikator wird insbesondere auch von der Zielgruppe akzeptiert und als sinnvoll empfunden.

Es ist oft nötig, mehrere Indikatoren für ein Ziel zu definieren. In der Praxis werden quantitative und qualitative Indikatoren häufig kombiniert. Grundsätzlich gilt aber, dass man sich auf ein Minimum beschränken sollte, um keine unnötigen Datenmengen zu produzieren.

Zu einem Indikator gehört immer auch die Angabe, aus welcher Quelle die Daten zum Indikator stammen und mit welcher Methode diese erhoben werden. Dadurch wird sichergestellt, dass der Indikator messbar ist. Zeigt es sich dabei, dass die Daten zum Indikator nicht oder nur mit unverhältnismässigem Aufwand erhoben werden können, muss der Indikator durch einen einfacheren ersetzt werden. Zu berücksichtigen ist auch die Möglichkeit, auf bestehende Quellen zurückzugreifen, z.B. auf nationale Statistiken oder auf Daten von Partnerorganisationen.

ACHTUNG

In der Praxis wird den Quellen oft zu wenig Aufmerksamkeit gewidmet und die Daten sind später gar nicht verfügbar, was den Indikator nutzlos macht. Ein Indikator ohne echte Datenquelle ist nicht messbar und kann damit nicht für eine Wirkungsmessung verwendet werden.

In der Praxis gibt es verschiedene Varianten Ziel und Indikator zu definieren bzw. voneinander zu unterscheiden. Diese sind gleichberechtigt. Man sollte sich jedoch immer bewusst sein, welche Definition gilt und diese konsistent anwenden. Diskussionen darüber sind fast unvermeidbar. Zu beachten ist, dass ein Indikator im hier verwendeten Sinne (der Indikator gibt an, ob das Ziel erreicht wurde oder nicht) immer auch eine Zielsetzung und damit die Wahl eines Vergleiches enthält.

Indikatoren sollen nicht nur auf den Wirkungsebenen (Outcome), sondern auch auf der Leistungsebene (Output) festgelegt werden.

Beispiele

- **Programm Gesundheit Kinder**

Ziel: Der Gesundheitszustand der Kinder in den ärmsten Regionen des Landes soll verbessert werden.

Indikator: Bis im Jahr 2015 ist die Kindersterblichkeit in den Regionen x, y, z auf % reduziert.

- **Projekt Brunnenbau**

Ziel: Die Bevölkerung in den Regionen x, y und z soll besseren Zugang zu sauberem Trinkwasser haben.

Indikator: Bis im Jahr 2015 haben 80% der Bevölkerung in den Regionen x, y und z innerhalb von 15 Minuten Zugang zu sauberem Trinkwasser.

- **Programm Bildung**

Ziel: Die Gemeinschaften haben besseren Zugang zu formeller und informeller Bildung

Indikator: 70% der 14'000 Personen, die neu Lesen und Schreiben können, bestätigen, dass ihre Lebensgrundlage dadurch verbessert wird.

Projekt Empowerment

Ziel: Sozial und ökonomisch Benachteiligte beeinflussen die Entscheidungsfindung in der Region

Indikator: Für 90% der lokalen politischen Vorhaben wurden öffentliche Anhörungen durchgeführt.

- **Projekt Förderung von Bauernorganisationen**

Ziel: Die Bauernorganisationen verbessern ihre institutionellen und organisatorischen Fähigkeiten.

Indikator: 100% der Bauernorganisationen bezeichnen in der Selbstevaluation ihre institutionellen und organisatorischen Fähigkeiten als mittel oder gut.

Ziel: Die Bauernorganisationen verbessern das Management ihrer ökonomischen Aktivitäten.

Indikator: 60% der Bauernorganisationen bezeichnen in der Selbstevaluation ihre Leistungen bezüglich der Entwicklung eines Partnernetzwerks als mittel oder gut.



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beispiel Projekt Gesundheitskurs

Für Mütter mit kleinen Kindern wird ein Kurs angeboten, der sie über den Zusammenhang von sauberem Trinkwasser und Gesundheit aufklärt.

	Ziel	Indikator	Messung
Wirkungsziel (W1)	Teilnehmerinnen kennen den Zusammenhang von sauberem Trinkwasser und Gesundheit.	Die Teilnehmerinnen können die vermittelten Inhalte im Rahmen eines Rollenspiels anwenden.	Video, Auswertung durch Projektverantwortliche
Leistungsziel (L1)	Kurse durchführen	100 Kurse	Projektbericht
Leistungsziel (L2)	Die Kurse sind gut besucht.	Mindestens 35 Teilnehmerinnen pro Kurs	Teilnehmerlisten



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beispiel Projekt Brunnenbau

Durch den Bau und die Reparatur von Brunnen wird der Zugang der Bevölkerung zu sauberem Trinkwasser verbessert.

	Ziel	Indikator	Messung
Wirkungsziel (W1)	Verbesserter Zugang zu sauberem Trinkwasser	Die Gehzeit zum nächsten Brunnen ist <15 Minuten für 80% der Haushalte	Beobachtung
Leistungsziel (L1)	Neue Brunnen bauen	50 neue Brunnen in der Region	Liste
Leistungsziel (L2)	Defekte Brunnen reparieren	80 Brunnen repariert	Liste



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beispiel Logframe Projekt

Gesundheitskurse

	Interventionsstrategie	Indikator	Quelle
Impact	Trägt dazu bei, dass Mütter vermehrt sauberes Wasser verwenden	90% der Haushalte verwenden hauptsächlich Wasser aus sauberen Quellen.	Umfrage
Outcome	Mütter kennen Zusammenhänge	Teilnehmerinnen der Kurse können die vermittelten Inhalte im Rahmen eines Rollenspiels anwenden.	Video, Auswertung durch Projektverantwortliche
Output	Mütter besuchen Kurse	100 Kurse durchgeführt, durchschnittlich über 35 Teilnehmerinnen pro Kurs	Projektbericht Teilnehmerlisten
Aktivitäten	Kurse für Mütter durchführen		



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beispiel Projekt medizinische Versorgung

Durch das Einrichten mobiler Krankenstationen können an Durchfall erkrankte Kinder medizinisch behandelt werden.

	Ziel	Indikator	Messung
Wirkungsziel (W1)	An Durchfall erkrankte Kinder können öfter geheilt werden.	95% der Behandlungen von Kindern mit Durchfall sind erfolgreich.	Fallstudien
Leistungsziel (L1)	An Durchfall erkrankte Kinder werden medizinisch behandelt.	1000 Behandlungen von Kindern mit Durchfall pro Jahr	Behandlungsstatistik
Leistungsziel (L2)	Mobile Krankenstationen kommen regelmässig in die Region.	3 mobile Stationen im Einsatz	Fahrplan, Einsatzplan



Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Erhebungsmethoden

Im Rahmen der Wirkungsmessung können verschiedene Erhebungsmethoden angewendet werden. Grundsätzlich kann zwischen qualitativen und quantitativen Erhebungsmethoden unterschieden werden. Neben der Wahl der Erhebungsmethode gilt es, in einem zweiten Schritt zu überlegen, ob alle oder nur ein Teil der betroffenen Einheiten, Zielgruppen oder Fälle in der Erhebung zur Wirkungsmessung berücksichtigt werden sollen (Untersuchungseinheit). Weiter gilt es, die Qualität der erhobenen Daten kritisch zu prüfen (Datenqualität).

Methodenmix

Im heutigen Forschungsalltag ist es üblich, qualitative und quantitative Methoden kombiniert einzusetzen und so von den Vorteilen beider Methoden zu profitieren. Man spricht in diesem Fall von einem Methodenmix oder der Triangulation. Der Nutzen der Kombination von qualitativen und quantitativen Methoden gilt als unbestritten und ist zum Standard aussagekräftiger Wirkungsmessungen geworden. So kann z.B. bei der Messung der Effektivität eines Programms einerseits ein standardisierter Fragebogen an die Zielgruppen verteilt und andererseits ein Interview mit Mitarbeitenden oder eine Gruppendiskussion mit Expertinnen und Experten durchgeführt werden. In Abhängigkeit der gewählten Erhebungsmethode haben die erhobenen Daten eine bestimmte Form (Interviewprotokolle, Protokolle von Expertengesprächen, Prozentzahlen aus einer Befragung, Häufigkeiten einer Beobachtung usw.) und müssen daher mit geeigneten Auswertungsverfahren analysiert werden.

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Beispiel Logframe Programm

Gesundheit Kinder

	Interventionsstrategie	Indikator	Quelle
Impact	Trägt zu verbessertem Gesundheitszustand der Kinder bei: Kindersterblichkeit sinkt	Kindersterblichkeit in den Regionen x, y, z von 10% auf 2% reduziert	Nationale Statistik
Outcome	1. Mütter kennen Zusammenhänge	Teilnehmerinnen der Kurse können die vermittelten Inhalte im Rahmen eines Rollenspiels anwenden.	Video, Auswertung durch Projektverantwortliche
	2. Erkrankte Kinder können erfolgreich behandelt werden.	95% der Behandlungen von Kindern mit Durchfall sind erfolgreich.	Fallstudien
	3. Verbessertes Zugang zu sauberem Trinkwasser	Gehzeit zum nächsten Brunnen <15 Minuten für 80% der Haushalte	Beobachtung
Output	1. Kurse	100 Kurse durchgeführt	Projektbericht
	2. Behandlungen	1000 Behandlungen/Jahr	Behandlungsstatistik
	3. Brunnen	50 neue Brunnen in der Region	Projektbericht
Aktivitäten	1. Kurse durchführen		
	2. Mobile Krankenstation betreiben		
	3. Brunnen bauen		

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Quantitative Methoden

Bei der quantitativen Methodik geht es darum, Verhalten oder Veränderungen in Form von zahlenmässigen Ausprägungen möglichst genau zu beschreiben und zu erfassen.

Anzahl der Untersuchungseinheiten:	Viele
Annahmen:	Klare Vorstellung über relevante Zusammenhänge
Ausgangspunkt:	Überprüfen von Vorstellungen
Fokus:	Wissen der Forschenden steht im Zentrum
Absicht:	Theorieevaluation

Eignung

Quantitative Verfahren eignen sich mit ihrer standardisierten Befragungs- und Beobachtungsform für die Untersuchung grosser Stichproben und für die Anwendung statistischer Auswertungsverfahren zur objektiven Messung und Quantifizierung von Sachverhalten. Sie sind ideal, um objektive Daten über die Zeit zu vergleichen und daraus Entwicklungen abzulesen. Quantitative Datenerhebungsmethoden erlauben es, eine grosse Menge von Informationen über zuvor definierte Verfahren zu untersuchen. Die gewonnenen Informationen können mit Hilfe statistischer Verfahren und Auswertungstechniken analysiert und verglichen werden.

Erhebung

Quantitative Daten werden mit folgenden Verfahren erhoben:

- Strukturierte Beobachtung, Messung, Zählung
- Analyse von Sekundärdaten (Statistiken, Prozessdaten)
- Verschiedene Formen von Befragungen und Experimenten

Stichprobengrösse

Die zu wählende Stichprobengrösse ist davon abhängig, wie genau die aus der Befragung resultierenden Ergebnisse sein müssen. Am einfachsten ist es daher, wenn alle Untersuchungseinheiten befragt werden können. Man spricht dann von einer Vollerhebung. Bei einer Vollerhebung erübrigen sich statistische Tests zur Signifikanz von Unterschieden, weil die Daten nicht auf einer Sichtprobe basieren, welche auf die Grundgesamtheit hochgerechnet werden. Aus der folgenden Tabelle lässt sich ableiten, dass bei Untersuchungseinheiten, die weniger als 300 Fälle umfassen am besten eine Vollerhebung durchgeführt wird. Evident wird auch, dass mit 300 befragten Einheiten relativ verlässliche Aussagen zu grossen Grundgesamtheiten gemacht werden können.

N =Grösse der Grundgesamtheit	n =	n =
Grösse der Grundgesamtheit	minimale Stichprobengrösse bei einem Fehler von +/-3 Prozentpunkten	minimale Stichprobengrösse bei einem Fehler von +/-5 Prozentpunkten
100	92	80
200	169	132
300	234	169
400	291	196
500	341	217
1000	516	278
5000	880	357
10 000	964	370
100 000	1056	383
1 000 000	1066	384

Auswertung

Bei quantitativen Erhebungsverfahren erfolgt die Auswertung mittels verschiedener statistischer Verfahren und Kennzahlen wie Häufigkeiten, prozentuale Anteile oder Mittelwerte sowie mit komplexeren statistischen Verfahren.

Vorteile	Nachteile
Exakt quantifizierbare Ergebnisse	Keine Flexibilität während der Untersuchung durch die Standardisierung der Untersuchungssituation. Die Fragen sind schon vorher festgelegt, es ist kein individuelles Eingehen auf die Testpersonen möglich.
Ermittlung von statistischen Zusammenhängen möglich	Man ermittelt nicht die Ursachen für einen Befund oder eine Einstellung wie beispielsweise Unzufriedenheit. Zur Verringerung dieses Problems empfiehlt sich der Einsatz offener Fragen.
Möglichkeit, eine grosse Stichprobe zu untersuchen und damit repräsentative Ergebnisse zu erhalten	Man erhält keine Verbesserungsvorschläge. Dieser Nachteil kann ebenfalls durch die Integration offener Fragen verringert werden.
Hohe externe Validität durch grosse Stichprobe	
Grössere Objektivität und Vergleichbarkeit der Ergebnisse	

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme



Qualitative Methoden

Bei der qualitativen Methodik geht es um das Beschreiben, Interpretieren und Verstehen von Zusammenhängen.

Anzahl der Untersuchungseinheiten:	Wenige
Annahmen:	Wenig gefestigtes Wissen über Wirkungszusammenhänge
Ausgangspunkt:	Bedarf an vertieften Informationen
Fokus:	Wissen der Akteure/Akteurinnen steht im Zentrum
Absicht:	Theoriekonstruktion

Eignung

Qualitative Datenerhebungsmethoden erlauben es, einen bestimmten Untersuchungsgegenstand im Detail und in der Tiefe zu untersuchen. Dabei können auch neue und nicht erwartete Informationen hervorgebracht werden. Dies kann zu einem tiefen Verständnis des Untersuchungsgegenstandes führen, erschwert aber Verallgemeinerungen über den konkret betrachteten Gegenstand hinaus. Die qualitative Befragung bzw. Beobachtung zeichnet sich durch eine unverzerrte und umfassende Informationen liefernde Herangehensweise aus und ist hierdurch überall dort geeignet, wo man eine differenzierte und ausführliche Beschreibung individueller Meinungen und Eindrücke benötigt. Insbesondere zur Sammlung von detaillierten Verbesserungsvorschlägen und zur Erkundung von Ursachen (für Sachverhalte wie beispielsweise Unzufriedenheit) sind qualitative Methoden ideal.

Erhebung

Qualitative Daten werden mit folgenden Verfahren erhoben:

- Verschiedene Formen von Interviews (Einzelgespräche, Gruppeninterviews, Fokusgruppen)
- Dokumentenanalyse

Stichprobengrösse

Über die Anzahl der Gespräche, die geführt werden sollen, gibt es in der Literatur keine einheitliche Meinung. Die Vorstellung von einer geeigneten Stichprobengrösse bewegt sich zwischen einigen wenigen Gesprächen und rund 200 Personen, wobei in Abhängigkeit der untersuchten Fragestellung ab einer gewissen Anzahl Gespräche eine theoretische Sättigung eintritt. Das heisst, dass sich der Erkenntnisgewinn durch zusätzliche Gespräche nicht weiter steigern lässt. Die erforderliche Stichprobengrösse ist im Allgemeinen deutlich geringer als bei quantitativen Verfahren. Bei der Zusammensetzung der Stichprobe gelten die Grundsätze des theoretical sampling, das heisst, die Stichprobe sollte den theoretischen Überlegungen und der Fragestellung angepasst werden, heterogen zusammengesetzt sein und möglichst typische Vertreter enthalten.

Auswertung

Bei qualitativen Erhebungsverfahren erfolgt die Auswertung anhand verschiedener Formen der Inhaltsanalyse. Diese Formen gehen von der Zusammenfassung und der schrittweisen Reduktion des Datenmaterials aus. Wichtig: Auch qualitativ gewonnene Daten können quantitativ ausgewertet werden.

Vorteile	Nachteile
Flexible Anwendung der Methode, die Methode passt sich an den Untersuchungsgegenstand an und nicht umgekehrt.	Die Anforderungen an die Qualifikation der beobachtenden oder interviewenden Person sind recht hoch, die Qualität der Daten ist zu einem gewissen Teil auch davon abhängig.
Die Offenheit des Vorgehens ermöglicht es, neue, bisher unbekannte Sachverhalte zu entdecken.	Die Auswertung ist vor allem im Vergleich zu den quantitativen Methoden relativ aufwändig.
Da die Teilnehmenden keinerlei Vorgaben haben, erhält man eher wahre und vollständige Informationen über die subjektive Sicht der Gesprächspartner.	Aus qualitativen Daten kann man keine zahlenmässigen Mengenangaben ableiten.
Der Fokus wird von den Teilnehmenden selbst bestimmt, dadurch liegt er vor allem auf den für sie relevanten Sachverhalten.	
Durch die persönliche Interaktion gibt es die Möglichkeit, Hintergründe zu erfragen und Unklarheiten zu beseitigen.	
Hohe inhaltliche Validität durch nicht prädefinierte Vorgehensweise	
Tieferer Informationsgehalt durch offene Befragung	
Größere Subjektivität der Ergebnisse	

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme

Untersuchungseinheit

Die Auswahl der für die Wirkungsmessung zu untersuchenden Einheiten, Zielgruppen oder Fälle hängt maßgeblich vom Design bzw. den Vergleichen ab.

Idealerweise werden alle betroffenen Einheiten, Zielgruppen oder Fälle in der Erhebung zur Wirkungsmessung berücksichtigt. Man spricht in diesem Fall von einer Vollerhebung. In der Praxis sind Vollerhebungen aus spezifischen Gründen oder wegen den Kosten nicht immer möglich. Daher muss entschieden werden, welche Untersuchungseinheiten bei der Wirkungsmessung berücksichtigt werden sollen. Bei quantitativen Verfahren man in solchen Fällen von einer Stichprobe.

Stichproben haben gegenüber den Vollerhebungen den Nachteil, dass Informationen immer nur für einen Teil interessierender Beobachtungen erhoben wird. Entsprechend muss man sich überlegen, ob die Ergebnisse auf einer Stichprobenauswahl basieren, auch für die gesamte Untersuchungseinheit gelten. Ist dies nicht der Fall, dann wurde die Stichprobe nicht richtig gezogen bzw. die zu analysierenden Einheiten nicht korrekt ausgewählt.

Auswahlkriterien bei qualitativen Verfahren

Nicht nur bei quantitativen Verfahren stellt sich die Frage nach der Auswahl der Stichprobe. Auch bei qualitativen Verfahren muss man sich überlegen, welche Einheiten für die Untersuchung berücksichtigt werden sollen. Aus den Auswahlkriterien ergibt sich meistens automatisch die Anzahl der zu berücksichtigenden Einheiten. Oftmals werden pro Auswahlkriterium mindestens eine oder zwei Einheiten berücksichtigt. Aus theoretischer Sicht ist die Anzahl der Untersuchungen genügend, wenn das Prinzip der Sättigung eintritt. Als gesättigt gilt eine Stichprobe, wenn zusätzlich Erhebungen keine neuen Informationen mehr erbringen und mit dem erhobenen Material der Erkenntnisgewinn gesättigt ist. Zur Bestimmung der richtigen Auswahl der Analyseeinheiten kann das Vorgehen in drei Schritten angewandt werden:

- In einem ersten Schritt empfiehlt es sich, das inhaltliche Interesse an bestimmten Gruppen zu präzisieren.
- In einem zweiten Schritt kann darauf geachtet werden, ob alle möglichen Ausprägungen und Merkmale der zu untersuchenden Einheit bei der Auswahl berücksichtigt worden sind.
- Im dritten Schritt ist nach der Datenerhebung nochmals zu prüfen, welche Konstellationen und Merkmale nicht in der bisherigen Erhebung vorkommen. Gegebenenfalls sind zusätzliche Untersuchungseinheiten zu berücksichtigen. Das bedeutet auch, dass die Auswahl der Untersuchungseinheiten nicht wie bei den quantitativen Verfahren anhand eines einzigen Schrittes durchgeführt werden muss.

Auswahlkriterien bei quantitativen Verfahren

Um Fehler oder Verzerrungen aufgrund einer falschen Auswahl von Fällen zu vermeiden, sollte genau geklärt werden, welche Kriterien bei der Auswahl zu berücksichtigen sind.

werden wer oder was zur Population gehört, die untersucht werden soll: Es ist insbesondere darauf zu achten auch schlecht erreichbare (z.B. räumlich) und marginalisierte Gruppen wie ethnische oder religiöse Minderheiten oder Frauen bei der Stichprobenziehung angemessen berücksichtigt werden. Weiter muss die Grösse der Stichprobe bestimmt werden: Zentrales Kriterium ist dabei, wie genau die Ergebnisse sein müssen. Die Grösse der Population hat, zumindest bei eher grossen Populationen, wenig Einfluss auf die Mindestgrösse der Stichprobe (siehe dazu auch quantitative Methoden). Natürlich spielen in der Praxis dabei auch die verfügbare Zeit und Kosten eine wichtige Rolle.

Zur Auswahl von Stichproben bei quantitativen Verfahren gibt es verschiedene Möglichkeiten. Grundsätzlich zwischen «Zufallsstichproben» und «Nicht-Zufallsstichproben», welche nach bestimmten Kriterien zusammengefasst werden, unterschieden. Wenn die Stichprobe nach dem Zufallsprinzip zusammengesetzt sein soll, hat jede der Grundgesamtheit angehört, dieselbe Wahrscheinlichkeit als Stichprobenmitglied «gezogen» zu werden. Folgendes werden einige der wichtigsten Auswahlverfahren kurz beschrieben.

Zufallsstichproben

- **Einfache Zufallsstichprobe**

Es gelangt jede Einheit der Grundgesamtheit mit der selben Wahrscheinlichkeit in die Auswahl (z.B. Namen aus einer Urne ziehen oder jedes x-te Haus).

- **Geschichtete Zufallsstichprobe**

Die Untersuchungseinheiten werden nach einem bestimmten Merkmal in Gruppen (Schichten) unterteilt (z.B. Dörfer, Kurse). Aus diesen Teilgesamtheiten werden dann nach dem Zufallsprinzip Stichproben entnommen.

- **Gestufte Zufallsstichprobe**

Es werden zuerst die Stufungskriterien bestimmt (z.B. Regionen A bis D). Dann wird die Grundgesamtheit nach diesen aufgeteilt und eine zufällige Auswahl getroffen (z.B. Region B und D) und auf eine bestimmte Anzahl von Primäreinheiten begrenzt, die man untersucht (z.B. 10 Brunnen je Region). Die restlichen Teilgesamtheiten werden ignoriert. Aus den zufällig ausgewählten Primäreinheiten (z.B. 10 Brunnen) ermittelt man nun die Zufallsstichprobe der Merkmalsträger (je 20 Haushalte im Umkreis von 15 Min.). In den beiden Regionen werden je 200 Haushalte befragt, die dann zu einer Gesamtstichprobe zusammengeführt werden.

Nicht-Zufallsstichproben

- **Quoten-Stichprobe**

Zuerst werden die Elemente der Grundgesamtheit in Gruppen eingeteilt. Die Stichprobe ist nun so zu ziehen, dass dieses Gruppenverhältnis in der Stichprobe möglichst genau so aussieht wie in der Grundgesamtheit. Es wird also versucht, eine gewünschte Populationsstruktur in der Stichprobe nachzuahmen. Dazu werden den Interviewern Vorgaben gemacht, welche Eigenschaften die zu Befragenden haben müssen. Wen er konkret auswählt, bleibt dabei dem Interviewer überlassen.

- **Homogene und heterogene Fallauswahl**

Die Beobachtungen werden so für die Stichprobe ausgewählt, dass sie möglichst ähnliche/unähnliche Eigenschaften aufweisen. In Fallstudien (welche per Definition keine Zufallsstichproben darstellen) werden beispielsweise häufig zwei Beobachtungen mit möglichst gegensätzlichen Eigenschaften untersucht.

- **Die Auswahl von typischen Fällen**

Hier werden diejenigen Beobachtungen für die Stichprobe ausgewählt, von denen man weiss oder annimmt, dass sie typische, durchschnittliche und eben keine extremen Eigenschaften aufweisen.

- **Die Auswahl von kritischen Fällen**

Es werden insbesondere oder ausschliesslich diejenigen Beobachtungen ausgewählt, von denen

Wirkungsmessung in der Entwicklungszusammenarbeit

Zewo-Leitfaden für Projekte und Programme

Datenqualität

Die Informations- oder Datenqualität bezeichnet die Relevanz und Korrektheit von Informationen. Sie gibt an, wie gut Daten die Realität oder die tatsächlichen Situationen beschreiben. Die Qualität der erhobenen Daten ist von zentraler Bedeutung dafür, dass eine Wirkungsmessung exakte Resultate liefern kann. In der wissenschaftlichen Methodik der Datenerhebung gibt es zwei Gütekriterien, welche die erhobenen Daten erfüllen sollten:

- **Reliabilität** (Zuverlässigkeit)
Der Begriff Reliabilität bezieht sich auf die Zuverlässigkeit der Datenerhebung. Ein Erhebungsverfahren ist dann zuverlässig, wenn unter den gleichen Bedingungen eine Wiederholung der Erhebung oder Messung zu den gleichen Ergebnissen führt.
- **Validität** (Gültigkeit)
Der Begriff Validität bezieht sich auf die Gültigkeit der Datenerhebung. Eine Datenerhebung ist dann gültig, wenn sie das misst, was man messen wollte. Eine Messung oder Erhebung ist valide, wenn die erhobenen Werte geeignete Kennzahlen für die zu untersuchende Fragestellung liefern.

Daten und Datenquellen prüfen

Daten und Datenquellen sollten bezüglich der Kriterien Reliabilität und Validität überprüft werden. Das ist insbesondere dann nötig, wenn es sich um fremde Datenquellen handelt oder wenn neue Datenquellen verwendet werden. Neue Erhebungs- und Verarbeitungsmethoden sollten ebenfalls getestet werden. Es kann sich als sinnvoll erweisen, die Datenerhebung vorgängig in einem Testlauf zu prüfen. Zusätzlich sollte geprüft werden, ob die Erhebung gewünschten Informationen liefert.

Fehlerquellen identifizieren und minimieren

Bei der Erhebung und der Erfassung von qualitativen und quantitativen Daten gibt es verschiedene Fehlerquellen, die vermieden werden sollten. Wird eine eigene Datenerhebung mittels qualitativer oder quantitativer Erhebungsinstrumente durchgeführt, dann sollten die Instrumente (Fragebogen, Gesprächsleitfaden usw.) vorab mit einer Pilotstudie (Pretest) geprüft werden. Dazu wird das Erhebungsinstrument an Testpersonen oder an Personen der Zielgruppe erprobt. Diese sollten, wenn möglich, der Zielgruppe der Befragung respektive den zu analysierenden Fällen entsprechen. Zudem sollte der Pretest unter möglichst ähnlichen Bedingungen durchgeführt werden wie die geplante Erhebung. In Abhängigkeit der Ergebnisse des Pretests ist eventuell eine Überarbeitung oder Anpassung der Erhebungsinstrumente notwendig. Es ist daher wichtig, dass die dazu benötigte Zeit bei der Planung berücksichtigt wird. Häufige Fehlerquellen liegen bei der korrekten Auswahl der Untersuchungseinheit.

man weiss, dass ihr Einschluss in die Stichprobe für die Glaubwürdigkeit oder Akzeptanz einer Studie ausschlaggebend ist.



Wirkungsmessung in der Entwicklungszusammenarbeit Zewo-Leitfaden für Projekte und Programme



Auftrag formulieren

Wenn die Wirkung durch ein internes Team gemessen und beurteilt wird, besonders aber wenn damit ein externes oder gemischtes Team beauftragt wird, sollte dieser Auftrag und die Planung der Wirkungsmessung schriftlich festgehalten werden.

Ein Auftrag für eine Wirkungsmessung kann der folgenden Struktur für Evaluationen folgen:

Muster zur Struktur für Evaluationen

Zweck und Nutzen (rationale, purpose)

Weshalb wird die Evaluation durchgeführt und wozu werden die Ergebnisse genutzt?

Zum Beispiel:

- Prozesse oder Verfahren optimieren
- Strategie oder Politik der Organisation weiterentwickeln
- Über die Weiterführung eines Projekts oder Programms entscheiden
- Rechenschaft gegenüber Geldgebern oder der Öffentlichkeit ablegen

Ziel (objectives)

Worüber soll die Evaluation etwas aussagen?

Zum Beispiel:

- Bestätigen, dass eine bestimmte Leistung (Output) erbracht, eine bestimmte Wirkung bei der Zielgruppe (Outcome) erreicht oder ein Beitrag zum übergeordneten Ziel geleistet wurde.
- Beurteilen, ob eine bestimmte Intervention effizient, effektiv und relevant war.
- Darlegen von Feststellungen, Schlussfolgerungen und Empfehlungen zu einem bestimmten Projekt oder Programm.

Umfang (scope)

Der Umfang der Evaluation soll klar definiert werden, indem die wichtigsten Eckpunkte klar umschrieben werden.

Zum Beispiel:

- Zu behandelnde Themen
- Zu betrachtende Zeitperiode
- Zu betrachtende Aktivitäten
- Bereits verwendete Mittel
- Geografische Ausdehnung
- Zielgruppen

Beteiligte und Betroffene

Wer ist an der Evaluation beteiligt oder von ihr betroffen?

Zum Beispiel:

- Projektverantwortliche Mitarbeitende
- Mittlerorganisationen
- Zielgruppe

- Partnerorganisationen
- Regierung

Welche Interessen und Bedürfnisse haben diese Personen(gruppen)? Wie werden diese berücksichtigt?

Berichterstattung

Wie wird über die Ergebnisse berichtet? Sind neben dem klassischen Bericht auch weitere «Produkte zu liefern»

Zum Beispiel:

- Klassischer Bericht
- Workshop mit den Beteiligten
- Debriefing mit Projektverantwortlichen
- Präsentationen für Linienverantwortliche
- Schriftliche «Lessons learnt»

Budget

- Kosten im Verhältnis zur Komplexität der Fragestellung und zum Wert der erwünschten Information verhältnismässig?

Was sollte schriftlich vereinbart werden?

- Ziel, Zweck, Blickwinkel der Evaluation festhalten
- Evaluationsfragen formulieren
- Termine und Budget vereinbaren
- Einem solchen Auftrag können Qualitätsstandards zugrunde gelegt werden (z.B. DAC , SEVAL, DEZA)
- Berichtsformat definieren



Stiftung Zewo
Schweizerische Zertifizierungsstelle für gemeinnützige,
Spenden sammelnde Organisationen

Stiftung Zewo
Lägerstrasse 27
8037 Zürich
Telefon +41 (0)44 366 99 55

Wirkungsmessung in der Entwicklungszusammenarbeit Zewo-Leitfaden für Projekte und Programme



DAC Evaluation Quality Standards

Der Entwicklungsausschuss der OECD (DAC) hat einen Leitfaden für gute Praxis bei der Evaluierung in der Entwicklungszusammenarbeit herausgegeben. Die Standards sollen die Qualität von Evaluierungsprozessen und -produkten verbessern und die Zusammenarbeit erleichtern. Die Grundsätze wurden im internationalen Konsens erarbeitet. Sie sind nach den typischen Phasen einer Evaluation gegliedert und enthalten von der Beschreibung des Zwecks, des Blickwinkels und des Kontexts über die Planung, Gestaltung, Durchführung und Berichterstattung bis hin zum Lernen und Nutzen der Ergebnisse jene Aspekte, die für die Qualität von Evaluationen von Bedeutung sind.

Download

DAC Guidelines and Reference Series Quality Standards for Development Evaluation, 2010, vollständiger Text

Link

Development Co-operation Directorate (DCD-DAC)



Stiftung Zewo
Schweizerische Zertifizierungsstelle für gemeinnützige,
Spenden sammelnde Organisationen

Stiftung Zewo
Lägerstrasse 27
8037 Zürich
Telefon +41 (0)44 366 99 55

Wirkungsmessung in der Entwicklungszusammenarbeit Zewo-Leitfaden für Projekte und Programme



SEVAL Evaluationsstandards

Die Schweizerische Evaluationsgesellschaft (SEVAL) ist eine multidisziplinäre Organisation, die sich für die Qualität und Verbreitung von Evaluationen einsetzt. Sie bildet ein Forum für den Austausch über Evaluationen zwischen Politik, Verwaltung, Hochschulen, NGOs und der Beratung. Die Evaluationsstandards der SEVAL sollen einen Beitrag zur Professionalisierung der Evaluation in der Schweiz leisten. Sie definieren die Qualitätsanforderungen, die an Evaluationen gestellt werden und richten sich insbesondere an Evaluatorinnen und Evaluatoren sowie Auftraggeberinnen und Auftraggeber. Die Leitlinien sind für Evaluationen aller Art ausgelegt, mit Ausnahme von Personalevaluationen. Sie sind in vier Themenbereiche gegliedert: Die Nützlichkeitsstandards stellen sicher, dass sich eine Evaluation an den Informationsbedürfnissen der vorgesehenen Nutzerinnen und Nutzer ausrichtet. Die Durchführbarkeitsstandards stellen sicher, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst ausgeführt wird. Die Korrektheitsstandards stellen sicher, dass eine Evaluation rechtlich und ethisch korrekt durchgeführt wird und dem Wohlergehen der Beteiligten und Betroffenen gebührende Aufmerksamkeit widmet. Die Genauigkeitsstandards stellen schliesslich sicher, dass eine Evaluation gültige und verwendbare Informationen hervorbringt und vermittelt.

Download

SEVAL Standards, Evaluationsstandards der Schweizerischen Evaluationsgesellschaft, 2000, vollständiger Text

Link

SEVAL



Wirkungsmessung in der Entwicklungszusammenarbeit Zewo-Leitfaden für Projekte und Programme



Leitprinzipien zu DEZA-Evaluationen

Die Direktion für Entwicklung und Zusammenarbeit (DEZA) will jährlich **0,6 bis 0,8 Prozent** ihres Gesamtbudgets für Evaluationen und Reviews aufwenden, um professionell über ihre Tätigkeit Rechenschaft abzulegen. Damit liegt sie im Durchschnitt vergleichbarer Agenturen. Bei bilateralen Gebern bewegen sich die Budgets für Evaluationen zwischen **0,1 und 2,5 Prozent** der Totalausgaben. Die DEZA publiziert das jährliche Evaluationsprogramm und die abgeschlossenen Evaluationen auf ihrer Website.

Mit ihrer Evaluationspolitik klärt die DEZA den Stellenwert der Evaluation und ordnet diese in einen grösseren Kontext ein. Sie zeigt die wichtigsten nationalen und internationalen Trends im Bereich der Evaluation auf und gibt eine Übersicht über ihre Evaluationsarchitektur. Die Politik hat normativen Charakter und dient als Grundlage für die Festlegung der minimalen Evaluationsstandards zur Qualitätssicherung. Auf eine vertiefte Diskussion einzelner Evaluationsmethoden wird dabei bewusst verzichtet. Die DAC/OECD-Evaluationsstandards, die Standards für die humanitäre Hilfe (ALNAP-Standards) und die Standards der Schweizerischen Gesellschaft für Evaluation (SEVAL) bilden den verbindlichen Rahmen.

In der Evaluationspolitik der DEZA sind 10 Leitprinzipien festgehalten, die die Grundwerte der DEZA-Evaluationstätigkeit reflektieren und den übergeordneten verbindlichen Rahmen für alle Mitarbeitenden bilden. Sie umfassen die folgenden Themenbereiche: Unabhängigkeit der Evaluationsteams, Unparteilichkeit, Objektivität und Glaubwürdigkeit, Transparenz, Partnerschaft, Machbarkeit, Nützlichkeit, Komplementarität, Subsidiarität, Steuerungsrelevanz und Persönlichkeitsschutz.

Downloads

DEZA-Evaluationspolitik, Direktion für Entwicklung und Zusammenarbeit, 2008, vollständiger Text
ALNAP-Standards, Evaluating humanitarian action using the OECD-DAC criteria, 2006

Links

DEZA
ALNAP