

Lernen aus Evaluierungen

Meta-Evaluation & Evaluationssynthese von InWEnt-Abschlussevaluierungen 2009

– Zusammenfassender Kurzbericht –

von

Prof. Dr. Alexandra Caspari

Fachhochschule Frankfurt am Main -
University of Applied Sciences



Fb4, Soziale Arbeit und Gesundheit
Nibelungenplatz 1,
D60318 Frankfurt am Main
+49 (0) 69 / 1533 - 2660
caspari@fb4.fh-frankfurt.de
www.fh-frankfurt.de/caspari

September 2010

1 Ziel, Nutzer/innen und Herangehensweise

1.1 Ziel und Nutzer/innen der Studie

Die Studie „*Meta-Evaluation & Evaluationssynthese von InWEnt-Abschlussevaluierungen 2009*“ ist im Auftrag von InWEnt – Internationale Weiterbildung und Entwicklung gGmbH, Abteilung 1.04 Qualität und Evaluierung, jetzt Abteilung Qualität, Wirkung und Nachhaltigkeit des Bereichs 1 Planung und Steuerung entstanden. Mit der Reorganisation der Evaluationsaufgaben bei InWEnt im Februar 2010 ist die Thematik in den Zuständigkeitsbereich der neu implementierten Evaluierungsbeauftragten 1.00 des Bereichs 1 Planung und Steuerung übergegangen.

Von Seiten InWEnt wurden die Berichte von allen in 2009 fertig gestellten bzw. der damaligen Abteilung 1.04 Qualität und Evaluierung übermittelten Abschlussevaluierungen zur Verfügung gestellt. Insgesamt liegen der vorliegenden Analyse *15 Evaluationsberichte* zugrunde – überwiegend Abschlussevaluierungen von in 2008 beendeten Programmen.

Diese 15 übermittelten Evaluationsberichte wurden einer unabhängigen Querschnittsauswertung unterzogen: Hierbei wurde einerseits im Rahmen einer *Meta-Evaluation* die *Qualität* der Evaluationen untersucht und andererseits mittels einer *Evaluationssynthese* die in den Berichten aufgeführten *Informationen, Schlussfolgerungen und Empfehlungen* in Hinblick auf übergreifende Erkenntnisse analysiert. Ziel der Meta-Evaluation war zu untersuchen, inwieweit die vorliegenden *Evaluationsberichte* bzw. *Evaluationsergebnisse methodisch zuverlässig* sind, um auf Basis fundierter Erkenntnisse über zentrale Problemfelder bei der Umsetzung der Evaluationen schlüssig und entscheidungsvorbereitend Verbesserungspotentiale für die Planung und Durchführung zukünftiger Abschlussevaluierungen aufzuzeigen. Weiteres Ziel war die Identifikation von Good Practices im Sinne methodisch guter Evaluierungen, um diese im Intranet zur Verfügung zu stellen. Die Evaluationssynthese wiederum hatte zum Ziel, die *inhaltlichen Ergebnisse* der Evaluationen im Querschnitt auszuwerten, um programmübergreifende Erfolgs- und Misserfolgskriterien systematisch herauszuarbeiten und Good Practice Beispiele zu identifizieren. Die Erkenntnisse der Studie sollen für die Weiterentwicklung der InWEnt Programme, Ansätze und Konzepte sowie der Instrumente und Verfahren des Programmintegrierten Planungs-, Monitorings- und Evaluierungssystems (PriME) genutzt werden.

Die Studie ist die *erste systematische Auswertung von InWEnt-Abschlussevaluierungen*. Sie richtet sich an die Evaluierungsbeauftragte, seit Februar 2010 u.a. zuständig für die konzeptionelle Weiterentwicklung von Evaluierungsinstrumenten, für die Beratung und Qualitätssicherung der dezentralen Evaluierungsvorhaben der programmführenden Abteilungen, für das zentrale Unabhängige Evaluierungsprogramm sowie für die Steuerung von verschiedenen Formaten des Lernens aus Evaluierungen. Die Studie richtet sich ebenso an die jetzige Abteilung 1.04 Qualität, Wirkung und Nachhaltigkeit im Rahmen des Qualitätsmanagements u.a. zuständig für die Qualitätssicherung von Planung, Monitoring und Evaluierung im Rahmen von PriME sowie die Weiterentwicklung der PriME-Instrumente, an die Abteilung 1.01 Programmkoordination und Ressourcenmanagement, zuständig für Antragstellung und Berichterstattung an das Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung

(BMZ), an die Abteilungen 1.02 und 1.05 Regionalkoordination, die Bereichs- und Abteilungsleiter/innen sowie die Programmverantwortlichen, die die Abschlussevaluierungen verantworten. Die vorliegende Studie ist im Kontext von PriME-Schritt 7 ‚Lernen aus Evaluierungen‘ zu sehen.

1.2 Methodik und Herangehensweise

Methodik der Meta-Evaluation

Für die *Meta-Evaluation* wurden auf Basis diverser Evaluations-Standards *Kriterien* festgelegt, anhand derer die Berichte ausgewertet wurden: neben grundlegenden deskriptiven Informationen wie Titel, beauftragende Abteilung, Sektor und Land der Maßnahme sowie Zeitpunkt der Evaluation wurden auch der Umfang der Berichte (Seitenzahl), die Sprache sowie die den Evaluatoren/innen zur Verfügung stehenden Arbeitstage erhoben. Die zentralen Kriterien der Meta-Evaluation wurden auf Basis bereits vorliegender Kriterienkataloge festgelegt: Grundlage waren das InWEnt ‚Standard-Berichtsraster für Evaluierungen‘ (PriME-Schritt 6.1) und die InWEnt ‚Checkliste für Abschlussevaluierungen‘ (PriME-Schritt 6.2), die sich stark an den Standards der DeGEval – Gesellschaft für Evaluation orientieren, sowie die erst als Entwurf vorliegenden ‚Qualitätsstandards für Evaluierungsberichte‘ des BMZ, welche zusätzlich zur Berücksichtigung der DeGEval-Standards auch dezidiert auf die Standards des Development Assistance Committee (DAC) Bezug nehmen. Diese Checklisten wurden an verschiedenen Stellen um spezifische Kriterien ergänzt, die der aktuellen internationalen Diskussion um adäquate Wirkungsmessung Rechnung tragen und insbesondere die genutzten Designs näher betrachten.¹

Die Kriterien beziehen sich einerseits auf die Bewertung der *Berichtsqualität*:

- 1) **Bericht allgemein:** z.B. inwieweit das Berichtsraster eingehalten wurde, der Bericht logisch aufgebaut bzw. klar gegliedert ist, die Zusammenfassung die zentralen Aussagen enthält, der Bericht insgesamt verständlich und nachvollziehbar ist.
- 2) **Darstellung der genutzten Methoden:** z.B. inwieweit die methodische Vorgehensweise transparent dargestellt wurde.

Andererseits wurden Kriterien zur *inhaltlichen Analyse* formuliert:

- 3) **Genutzte Methoden selbst:** inwieweit alle Personengruppen berücksichtigt wurden, die Daten aus InWEnts Planung, Monitoring und Evaluierung (PM+E) berücksichtigt wurden, die methodische Vorgehensweise dem Gegenstand angemessen war, die Evaluatoren/innen ausreichend neutral bzw. objektiv waren und zwischen Ergebnisse und Analyse bzw. Interpretation getrennt wurde. Vor dem Hintergrund der aktuellen Diskussion um adäquate Wirkungs-

¹ Die Kriterien entsprechen somit *aktuellen Anforderungen* an Evaluierungen und wurden größtenteils auf Basis der aktuell gültigen Version des PriME-Handbuchs von 2008 erarbeitet. Bei den aufgezeigten Ergebnissen insbesondere der Meta-Evaluation aber auch der Evaluationssynthese ist zu berücksichtigen, dass die analysierten Evaluierungsberichte Maßnahmen zum Gegenstand hatten, die teilweise *vor* der Einführung von PriME 2005 in seiner ersten Version geplant wurden. Bereits mit der Überarbeitung von PriME 2008 wurden umfangreiche verbindliche Änderungen vorgenommen. Des Weiteren wurden insbesondere in 2009/2010 einzelne PriME-Instrumente überarbeitet und erweitert.

untersuchungen bzw. Evaluationsdesigns wurde die methodische Vorgehensweise näher betrachtet: so wurde detailliert untersucht, inwieweit die Vorgehensweise systematisch war, d.h. ob eine Wirkungskette aufgezeigt wurde, Hypothesen und Indikatoren formuliert wurden und ob die theoretische Konzeption sowie das Untersuchungsdesign angemessen waren. Ein zweiter Unterpunkt zur Analyse der Angemessenheit der methodischen Vorgehensweise bezog sich dezidiert auf die genutzten Datenerhebungsmethoden – hier wird überprüft, welche Methoden eingesetzt wurden: Dokumentenanalyse von InWEnt-Unterlagen, Analysen von weiteren (wissenschaftlichen) Dokumenten, leitfadengestützte Intensivinterviews, standardisierte Befragungen, systematische Beobachtungen. Des Weiteren wurde überprüft, ob ein ausreichender Methoden-Mix zugrunde lag und inwieweit partizipative Konzepte genutzt wurden. Der dritte Unterpunkt zur Überprüfung einer angemessenen methodischen Vorgehensweise hinterfragt inwieweit das Kausalitätsproblem angemessen behandelt wurde: wurden Kontrollgruppen berücksichtigt, Vorher-Daten (Baseline) genutzt, ist die Stichprobenziehung angemessen.

4) Inhalte: inwieweit sind die Informationen ausreichend für die Begründung der Schlussfolgerungen und Empfehlungen, werden die unterschiedlichen Perspektiven der Beteiligten ausreichend dargestellt, werden die Stärken und Schwächen der evaluierten Gegenstandes ausreichend dargestellt, werden alle Ebenen behandelt (Individuum, Organisation, System), werden alle Ebenen behandelt (Input, Output, Outcome, Impact), werden die DAC-Kriterien bearbeitet (Relevanz, Effektivität, Impact, Effizienz, Nachhaltigkeit; sowie Kohärenz, Komplementarität und Koordination).

5) Schlussfolgerungen und Empfehlungen: inwieweit sind die Schlussfolgerungen aus den erhobenen Informationen gut begründet, wurde die Informationen angemessen analysiert, sind die Empfehlungen aus den erhobenen Informationen und Schlussfolgerungen abgeleitet.

Alle Evaluationsberichte wurden anhand der aufgeführten Kriterien analysiert. Die Qualitätsbeurteilung erfolgte je Kriterium entweder in Form einer Ziffer in vier Abstufungen 1 ‚sehr gut‘, 2 ‚eher gut‘, 3 ‚eher schlecht‘ und 4 ‚sehr schlecht‘ oder aber es wurde lediglich vermerkt, dass das Kriterium erfüllt wurde (,x‘). Zusätzlich wurden zu einigen Kriterien Begründungen für die Bewertung bzw. Erläuterungen notiert.

Da insbesondere bzgl. der unter Pkt. 1 aufgeführten Aspekte nach Auswertung der Ergebnisse viele Fragen offenblieben, wurden nachträglich auch die Terms of Reference (TORs) zu den jeweiligen Berichten mit Blick auf folgende Kriterien analysiert: Bezeichnung der Evaluation, geforderte Berichtssprache, geforderte Berichtsstruktur, geforderte Datenerhebungsmethoden, geforderter Bezug zu DAC-Kriterien sowie vorgegebene Fachkrafttage.

Methodik der Evaluationssynthese

Für die *Evaluationssynthese* wurde ebenfalls ein Analyseraster erstellt. Dieses umfasste z.B. den Zeitraum der Programmdurchführung, die Programmkosten, eine Kurzbeschreibung zum Programm, der Zielgruppe sowie den Zielen (Programmziele und Maßnahmenziele). Hauptaugenmerk der Evaluationssynthese richtete sich auf die in den Berichten enthaltenen *zentra-*

len Ergebnisse sowie zentralen Erfolgsfaktoren. Aus allen Einzelberichten wurden Textpassagen zu diesen Kriterien über Copy-Paste herausgefiltert. Im Ergebnis zeigte sich, dass die erhaltenen Informationen bzw. Textpassagen nicht ausreichend für eine inhaltliche Querschnittsauswertung sind. Die Meta-Evaluation verdeutlichte die Ursache: In den Berichten wurden die im Analyseraster aufgeführten Aspekte insbesondere die zentralen Erfolgsfaktoren in den wenigsten Fällen systematisch analysiert und gebündelt zusammengefasst in einem Kapitel dargestellt. Dennoch werden hinderliche Faktoren oder Gründe für Misserfolge über den Text verteilt benannt, teilweise in Form von Aufzählungen bzw. Darstellungen von Problemen oder von Seiten der Teilnehmer/innen geäußerter Verbesserungsbedarf. Daher wurden auf Basis der ersten Analyse alle in den einzelnen Berichten erwähnten Erfolgs- und Misserfolgskriterien systematisch herausgefiltert und thematisch sortiert. Anschließend wurden alle Berichte erneut im Hinblick auf Hinweise zu diesen Faktoren erneut analysiert.² Entsprechende Textpassagen wurden wiederum über Copy-Paste gesammelt. Hierdurch ergab sich eine *systematische Sammlung und Zusammenstellung* von diversen in den Evaluierungsberichten angeführten Erkenntnissen. Aus den mehrfach aufgeführten, identischen Aspekten konnten **programmübergreifende Erkenntnisse** zu den Wirkungen von InWEnt-Maßnahmen sowie offenkundigem *Verbesserungspotential* abgeleitet werden. Des Weiteren wurden in diversen Berichten ausführlich dargestellte oder besonders betonte **Einzelaspekte**, die offensichtlich entscheidend zum Erfolg beigetragen haben, herausgefiltert und aufgeführt, d.h. Best Practice Beispiele, die auch im Kontext andere InWEnt-Programme relevant erscheinen und auf ihre Übertragbarkeit geprüft werden sollten.

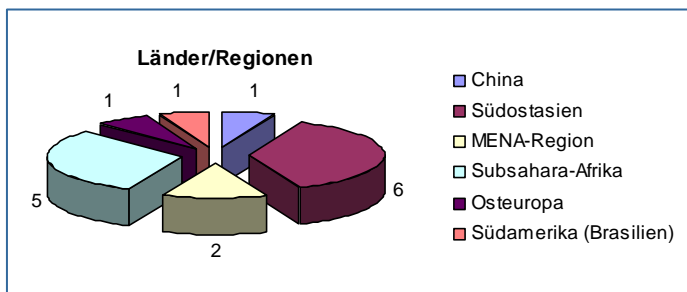
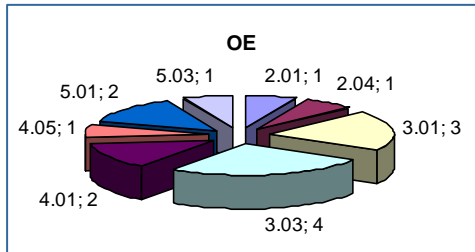
Die Evaluationssynthese entspricht insofern einer Textanalyse. Die diversen Aussagen aus den Berichten zu einzelnen identifizierten Aspekten werden gebündelt dargestellt und anhand von Zitaten belegt. Die Analyse greift hierbei wiederum Erkenntnisse aus der Meta-Evaluation auf, da hier mögliche Ursachen aber auch Ansätze zur Verbesserung liegen.

² Hierbei wurden auch zwei im Rahmen der Meta-Evaluation untersuchte Kriterien bzgl. des Berichtsrasters (Zusammenfassung und Schlussfolgerungen) erneut analysiert: zusätzlich zu der zuvor überprüften Frage, ob eine Zusammenfassung bzw. ein Kapitel mit Schlussfolgerungen vorhanden ist, wurden diese Kapitel bzw. Textpassagen nun auch *inhaltlich* betrachtet und überprüft, ob die für diese Kapitel relevanten Aspekte auch tatsächlich enthalten sind. Es wird an dieser Stelle deutlich, dass Meta-Evaluation und Evaluationssynthese ineinander übergehen und sich gegenseitig ergänzen.

2. Zentrale Erkenntnisse

2.1 Die Grundlage: Zahlen, Daten, Fakten

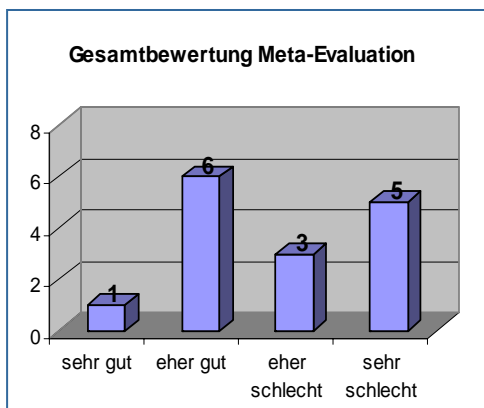
Die 15 der Studie zugrunde liegenden Berichte decken *alle InWent-Bereiche des Kerngeschäfts* ab: zwei Berichte wurden aus Bereich 2 ‚Internationale Regelwerke/Gute Regierungsführung/Wirtschaftspolitik‘ beauftragt, je drei aus Bereich 4 ‚Nachhaltiges Wirtschaften‘ sowie Bereich 5 ‚Umwelt, natürliche Ressourcen und Ernährung‘. Fast die Hälfte der Berichte (n:7) sind dem Bereich 3 ‚Soziale Entwicklung‘ zugeordnet, von denen wiederum vier dem Bereich ‚Gesundheit‘ (Abteilung 3.02) entstammen.



Ein Großteil der Evaluationen bezieht sich auf Länder/Regionen in Südostasien (n:6) sowie in Subsahara-Afrika (n:5). Sechs Berichte haben ein spezifisches Land zum Gegenstand, zwei beziehen sich auf zwei Länder, sieben auf mehrere Länder bzw. eine ganze Region.³

2.2 Lernen aus Evaluierungen – allgemein

Zentrales Ergebnis der *Meta-Evaluation* ist, dass in der zusammenfassenden Gesamtbetrachtung ein Evaluationsbericht als sehr gut bewertet, immerhin sechs weitere als eher gut werden, fünf Berichte dagegen sind keinesfalls ausreichend. Insgesamt ergibt die Meta-Evaluation ein mit internationalen Studien zur Qualität von EZ-Evaluationen vergleichbares Ergebnis⁴: Der Evaluationsprozess ist zu großen Teilen als *nicht transparent* (67%) und die genutzten Datenerhebungsmethoden meist als *inadäquat* (73%) zu werten, d.h. die *Evaluationsergebnisse und Erkenntnisse* insbesondere bzgl. der *Programmwirkungen* sind meist nicht zuverlässig bzw. nicht nachvollziehbar abgeleitet.



Zwar kann somit in den analysierten Berichten der *Nachweises von Wirkungen* aufgrund unzureichender Methoden nicht erbracht werden, allerdings ist hierbei zu beachten, dass die Notwendigkeit eines eindeutigen Wirkungsnachweises i.S.v. kausaler Attribution erst seit we-

³ Ein Bericht bezieht sich auf zwei Länder, daher summieren sich die aufgeführten Zahlen auf insgesamt 16.

⁴ vgl. Caspari, Alexandra, 2004: Evaluationen der Nachhaltigkeit von Entwicklungszusammenarbeit. Zur Notwendigkeit angemessener Konzepte und Methoden, Wiesbaden: VS-Verlag, S. 86ff.

nigen Jahren intensiv diskutiert wird⁵. Von daher sind entsprechende Evaluationsergebnisse erst künftig zu erwarten.

Zentrales Ergebnis der *Evaluationssynthese* ist, dass von den Evaluatoren/innen dennoch offensichtliche *zentrale Erkenntnisse* zu bedeutsamen bzw. markanten Programmervfahrungen sowie *offenkundigen Erfolgs- bzw. Misserfolgsk Faktoren* identifiziert und in den Berichten aufgezeigt werden. Aus den *mehrfach aufgeführten, identischen Aspekten* konnten somit *programmübergreifende Erkenntnisse* abgeleitet werden. Neben in mehreren Berichten wiederkehrenden Aspekten wurden auch in einigen Berichten ausführlich dargestellte oder besonders betonte *Einzelaspekte*, die auch im Kontext anderer Programme als relevant erscheinen, herausgefiltert und somit weitere *potentielle Best Practice Beispiele* identifiziert.



Meta-Evaluationen und Evaluationssynthesen führen zu programmübergreifenden systematischen Erkenntnissen. Alle *InWEnt-Abschlussevaluierungen* sollten daher in *jährliche Meta-Evaluationen und Evaluationssynthesen* einfließen. Des Weiteren ist zu prüfen, inwieweit auch eine regelmäßige Querschnittsauswertung der Zwischenevaluierungen sowie der Evaluierungen des Unabhängigen Evaluierungsprogramms möglich und sinnvoll ist.

2.3 PriME – Programmintegriertes Planungs-, Monitorings- und Evaluierungssystem

Zwar werden nicht immer alle PriME-Schritte, insbesondere PriME-Schritt 5 „Outcome-Monitoring“, im erforderlichen Maße umgesetzt, so zeigt die Analyse der 15 Berichte, dass die im Rahmen von PriME erarbeiteten *praktischen Arbeitshilfen* (z.B. die Fragebögen für Maßnahmenevaluierungen, Standard-Berichtsraster, Standard-TORs) in den Programmen umfangreich genutzt werden. Solche ausgearbeitete ‚Werkzeuge‘ scheinen für eine adäquate Umsetzung von PriME äußerst relevant.



Die im Rahmen von PriME erarbeiteten *praktischen Arbeitshilfen* werden in den Programmen umfangreich genutzt. Der bereits *begonnene Prozess zur Erweiterung und Verbesserung der Instrumente des Outcome-Monitorings* ist unbedingt weiter zu verfolgen. Es sind praxisbezogene Instrumente incl. weiterführender Standards und Arbeitshilfen zu entwickeln, die möglichst zeitnah in die Programmarbeit implementiert werden können.

Im Rahmen von PriME 2008 wurde ein InWEnt ‚*Standard-Berichtsraster für Evaluierungen*‘ (PriME-Schritt 6.1) entwickelt, das 2009 überarbeitet wurde. Dieses Standard-Berichtsraster wirkt sich positiv auf einen logisch gegliederten und verständlich formulierten Bericht aus. Es sollte daher nachdrücklich ein *Standard-Berichtsraster* für Evaluierungen mit Erläuterungen als verbindliche Vorgabe für Evaluationen gelten. Die Verbindlichkeit sowohl bzgl. der Gliederung sowie der Inhalte je Kapitel sollte stärker betont und bei der *Berichtsabnahme über-*

⁵ vgl.: Caspari, Alexandra, 2009: ‚Rigorese‘ Wirkungsevaluation – methodische und konzeptionelle Ansätze der Wirkungsmessung in der Entwicklungszusammenarbeit, in: Zeitschrift für Evaluation 8 (2), S. 183-213. sowie Caspari, Alexandra/Barbu, Ragnhild, 2008: Wirkungsevaluierungen. Zum Stand der internationalen Diskussion und dessen Relevanz für die Evaluierung der deutschen Entwicklungszusammenarbeit. Evaluation Working Papers. Bonn: BMZ.

prüft werden. Des Weiteren sollte das Berichtsraster sowie die **Standard-TORs** weiter überarbeitet werden, um den Programmen hiermit optimale Instrumente für Abschlussevaluierungen an die Hand zu geben. Es kann davon ausgegangen werden, dass sich dies positiv auf die Qualität der Abschlussevaluierungen und somit auch auf die Ergebnisse zukünftiger Meta-Evaluationen auswirkt.

Die Analyse der 15 Evaluationsberichte weist ferner darauf hin, dass die Programmleiter/innen bei der Umsetzung der gesamten PM+E-Schleife (Planung, Monitoring und Evaluierung) *intensive(re) Unterstützung* benötigen, insbesondere bei der Formulierung von Impact, Programm-/Maßnahmenzielen und angemessener Indikatoren, sowie bei der Implementierung und Umsetzung von M+E sowie dem Outcome-Monitoring. Hier sollte anhand einer Bedarfsanalyse eruiert werden, welche weiteren konkreten Unterstützungen notwendig und hilfreich wären (Themen für „Werkstatt-Tage“). Gleichzeitig sollte die Umsetzung aller PriME-Schritte insb. die Erstellung der M+E-Reports nachdrücklicher sichergestellt, d.h. verbindlicher eingefordert werden.



Die **Umsetzung von PriME** muss nachdrücklich sichergestellt werden. Weitere konkrete Unterstützung der Programmleiter/innen bei der Umsetzung der gesamten PM+E-Schleife scheint notwendig. Der verbindliche Charakter von PriME sollte über eine *Organisationsanweisung* erzeugt werden.

2.4 Erfolgsfaktoren für die Programmarbeit

Die *Vermittlung sektorübergreifender Schlüsselkompetenzen (Soft Skills)* scheint die **komparative Stärke von InWEnt** zu sein: In so gut wie allen Evaluierungsberichten werden positive Wirkungen der im Rahmen der Programme durchgeführten *Kurse* auf *Individualebene* aufgeführt. Die teilweise beachtlichen Wirkungen scheinen auf zwei zentrale Ursachen zurückzuführen sein: zum einen die mehrfach hervorgehobene *hohe Qualität* der InWEnt-Kurse allgemein und zum anderen spezifisch die im Rahmen der Kurse genutzten *didaktischen Methoden*. Die von InWEnt formulierten sieben didaktischen Prinzipien, insbesondere die hierbei hervorgehobene Bedeutung der Entwicklung von Selbstlernkompetenz sowie des antizipierenden Denkens, werden im Rahmen der InWEnt-Kurse optimal und beispielhaft umgesetzt und von den Teilnehmer/innen sehr geschätzt. Es scheint, dass die meisten der in den Evaluierungsberichten dargestellten positiven Wirkungen somit auf Veränderungen in den individuellen Handlungs- und Veränderungskompetenz basieren.



Auszug: Hohe Qualität der InWEnt-Trainingskurse

„In allen Kursen haben die Teilnehmenden viel **von der Methodologie selbst gelernt**; sie diente ihnen als Modell dafür, wie sie ihre eigenen Interventionen im Bereich Training und Moderation gestalten sollten“

„Die Erkundigung lässt keinen Zweifel, dass die meisten Teilnehmerinnen der beiden Kurse von den InWEnt **Kapazitätsverbesserungsmethodologie ganz begeistert waren** und manche fanden die neue Ausbildungsmethodologie am interessantesten von Allem was sie bei InWEnt gelernt haben“

“The design of the **learning process** exceeded the expectation of the participants. They expected a frontal teaching methodology as they are used to. Interactive learning was new to them and they appreciated their active involvement in the entire learning processes. Hence InWEnt **proved a client oriented design of the learning processes** and thereby achieved a high quality in their capacity programme”

...



...

„Als **komplementäre Stärken InWEnts** haben sich (...) die **Vermittlung interkulturellen Dialoges** (...) erwiesen“

“A large number of (...) participants **praised the professionalism and quality of the organisation, the right mix between practical and theoretical content** (e.g. inclusion of case studies and field trips), and InWEnt’s participatory approach to capacity building (e.g. learning from each other and course participant interactions)”



Die **Vermittlung von Soft Skills** scheint die **komparative Stärke von InWEnt** zu sein: Die hohe Qualität der Kurse (didaktische Methoden, Vermittlung sektorübergreifender Schlüsselqualifikationen) werden *sehr geschätzt* und *bewirken beachtliche Veränderungen auf individueller Ebene*.

Dagegen zeigen sich auf *organisatorischer, institutioneller Ebene* zum Zeitpunkt der Abschlussevaluierungen eher geringe Wirkungen. Die Analyse der Berichte weist auf diverse zentrale *Erfolgsfaktoren* hin, die für die Wirksamkeit auf institutioneller Ebene von entscheidender Relevanz sind:

Wenn auch die didaktischen Methoden der InWEnt-Kurse immer wieder positiv hervorgehoben werden, so scheint es gleichzeitig Verbesserungsbedarf bzgl. der **Kursinhalte** zu geben. In vielen Evaluierungsberichten wird darauf hingewiesen, dass die Angebote nicht immer den *lokalen, regionalen* aber auch *institutionellen und individuellen Anforderungen und Nachfragen* entsprochen haben.



Auszug: Inhalte von Trainingskursen auf Bedürfnisse ausrichten

„Participants (...) of the training courses noted (...) that **case studies lacked the Philippine context**“

“**Targeting of products** to training beneficiaries and job profiles: there is room for improving this by a rapid assessment of the participant’s educational background and professional experience, prior to the training (...) to create a differentiated learning approach **to better match the participants’ skills needs**“

„Das durch hohe Qualität gekennzeichnete regionale InWEnt-Programm könnte noch an Bedeutung gewinnen, wenn seine Inhalte mehr auf die **Bedürfnisse der einzelnen Länder** ausgerichtet wären“

„Für Afrika sollen die **Inhalte noch bessere auf den Bedarf fokussiert** werden“

„Verbesserungsbedarf besteht insbesondere bei der Planung von Ausbildungsangeboten unter Berücksichtigung **lokaler und regionaler Marktanforderungen und Nachfrage**“

“Most interventions do not **cater specific recipient needs** but rather **provide generic trainings on broad topics, disconnected from the actual capacity and performance of specific organisations**“

“Ensuring that the learning experience of the **individual participant contributes to organisational performance** improvements requires not only high-quality training but, in addition, careful analysis of specific **capacity building needs** (...)“

“key weakness of the entire programme is the **lack of specific identification of capacity building needs**“

Eine weitere Möglichkeit, Wirkungen auf institutioneller Ebene zu erhöhen, betrifft die **Auswahl der Kursteilnehmer/innen**. Aus einigen Berichten geht hervor, dass Teilnehmer/innen aus einem *homogenen Umfeld* ausgewählt wurden, was zu einer „kritischen Masse“ fortgebildeter Mitarbeiter/innen in einzelnen Institutionen führte. Aber auch eine homogenere Zusammensetzung der Kursteilnehmer/innen bzgl. *Vorkenntnisse und Fähigkeiten* scheint von Vorteil. Andere Berichte weisen darauf hin, dass verstärkt Teilnehmer/innen aus *höheren Managementebenen* mit mehr Entscheidungsbefugnissen innerhalb ihrer Organisationen bzw. *Entscheidungsträger aus Verwaltung oder Regierung* für InWEnt-Kurse gewonnen werden sollten.

**Auszug: Auswahl Teilnehmer/innen**

„a more **homogeneous group of participants** would benefit the implementation“

„Generell wird empfohlen, **Gruppe mit homogenen Fähigkeiten** zu bilden“

„Besonderes Gewicht sollte künftig auch auf die **verbesserte Auswahl von Teilnehmern** an InWEnt Förderungen gelegt werden“

„Die Konzentration auf das **Top Management** (...) stellen weitere Erfolgsfaktoren des Projekts dar“

„Die **Teilnehmenden** des Bildungskurses wurden aus einem **homogenen Umfeld rekrutiert**: (...) Dies hat zur Schaffung eines **Nodus fortgebildeter Personen, einer kritischen Masse**, in diesen Institutionen geführt. In Malawi sind diese Noduli nun im Begriff, systemische und strukturelle Veränderungen herbeizuführen. Dies steht im Gegensatz zum Gesundheits- und multisektoralen Kurs, in denen die Teilnehmenden aus sehr **heterogenen** Kontexten kommen“

„the **selection of suitable participants is of great importance**“

Ein weiterer Punkt, der in vielen Berichten angeführt ist, ist die Empfehlung die **Kooperation mit anderen EZ-Partnern sowie nationalen Counterparts** zu verstärken. Hierbei wird Kooperation auf verschiedenen Ebenen angesprochen: zum einen eine intensivere Kooperation mit der *Führungsebene der Partnerorganisationen* bei der *Auswahl der Teilnehmer/innen*, um deren Arbeitsumfeld zu optimieren und so die Chancen der institutionellen Wirksamkeit zu erhöhen. Insbesondere aber Kooperation im Sinne von *Abstimmung und Verzahnung geplanter Capacity Building Programme* der verschiedenen EZ-Organisationen, um Überschneidungen zu vermeiden und Synergien zu nutzen. Letztendlich wurde auch Kooperation im Sinne von *frühzeitiger Information über InWEnt-Capacity Building Programme* aufgeführt, denn teilweise kamen die Ankündigungen zu Weiterbildungskursen zu spät, um von Partnern in deren Programme integriert werden zu können – teilweise waren die InWEnt-Kurse aber auch Programme nicht bei allen Partnern bekannt:

**Kooperation, Abstimmung sowie Information***

„Eine **verstärkte Kooperation mit anderen Einrichtungen der deutschen EZ** in den Partnerländern sollte diese Trainingsbemühungen ergänzen“

„Gemäß von den Partnern der deutschen EZ (...) erhaltenen Informationen kommt die **Ankündigung der Trainingskurse zu spät**, als dass die Maßnahmen noch effektiv in die Programme der deutschen EZ integriert werden könnten (...). Auch wird der **Inhalt einiger Kurse** als nicht immer mit den auf **nationaler Ebene bestehenden Bedarfen** als ausreichend **abgestimmt** angesehen“

„Das derzeitige Programm als solches (...) ist **nicht bei allen Partnern der deutschen EZ bekannt**“

„Insbesondere wird empfohlen, die **Partnerinstitutionen der deutschen EZ in Deutschland als auch die Sektorbeauftragten und Länderdirektoren** mit angemessenem zeitlichem Vorlauf zu informieren, so dass sie die Trainingsmaßnahmen in ihre jährliche Planung integrieren können“

“The problem [activities ‘kidnapped’] could be mitigated by **aligning the time and cycle of planning between all partners** including the budgetary yearly versus long-term implications”

„Generell sollten alle **InWEnt Fortbildungsprojekte mit solchen TZ-Programmen anderer Geber verzahnt werden**, die auf eine institutionelle Konsolidierung der Partnerorganisation abzielen“

“(…) room for re-conceptualising (...): Improvements in coordination of staff development in general, and of Educational Planning and Budgeting **with other training activities carried out under the auspices of other government units to avoid overlaps and to create synergies**“

„(...) hätte eine **engere Kooperation** zwischen InWEnt und den Partnern der deutschen EZ (...) den **Auswahlprozess** in einigen Partnerländern optimieren können. Auch dürfte ferner eine **intensivierte Informationsstrategie** seitens InWEnt einer solchen Kooperation förderlich sein und Synergien für die Programmumsetzung zeitigen“

„Leider hat das Programm keine sehr enge Beziehungen mit den **verschiedenen Gesundheitsministerien** hergestellt, was auch für die wichtigsten Entwicklungspartner zutrifft“

“The course **benefited** from (...) and an **excellent relationship with the core ministries**“



Zentrale **Erfolgsfaktoren** für die Programmarbeit:

Auswahl der Kursteilnehmer/innen

- Es sollten verstärkt Personen aus *höherer Managementebene* mit mehr Entscheidungsbefugnisse für InWEnt-Kurse gewonnen werden
- Es sollte auf eine *homogenere Zusammensetzung* der Kursteilnehmer/innen geachtet werden (bzgl. Herkunftsorganisation sowie Vorkenntnisse und Fähigkeiten)

Bedarfsorientierung Kursinhalte

- Die Kursinhalte sind stärker auf *Bedarfe der Teilnehmer/innen und Partnerorganisationen* abzustimmen

Kooperation und Abstimmung

- mit *EZ-Partner & Stakeholder* vor Ort bei TN-Auswahl und Abstimmung der Kursinhalte
- mit *Bildungsanbieter und EZ-Partner* von Ort, um Fortbildungsprogramme optimaler zu Verzahnen und Abzustimmen sowie im Sinne einer frühzeitigen Information

In vielen Evaluationsberichten wurde ein weiterer offensichtlicher Erfolgsfaktor angesprochen, nämlich dass InWEnt-Kurse bzw. Programmmodule nur selten durch **Universitäten der Partnerländer akkreditiert** sind und somit für die Teilnahme **keine Credit Points** erlangt werden können, die im Rahmen einer weiteren Qualifikation genutzt werden können. Da die Bedeutung von Akkreditierung weltweit auch im Weiterbildungssektor mehr und mehr zu nimmt, sowohl auf der Angebots- als auch der Nachfrageseite, sollte InWEnt als international renommierte Weiterbildungsorganisation diese Entwicklung rechtzeitig berücksichtigen. Dass eine solche Akkreditierung einzelner InWEnt-Kurse *grundsätzlich möglich ist*, wird in einem Bericht aufgezeigt. InWEnt sollte daher unbedingt anstreben, zukünftig vermehrt Kurse von Hochschulen der Partnerländer akkreditieren zu lassen.



Auszug Akkreditierung*

„Strategische Bedeutung haben **Verankerung** der pädagogischen und fachlichen Ausbildung von Berufsschullehrer/innen **im Hochschulsystem**“

„Verbindung zu einer Partnerinstitution für die weitere Kursentwicklung und -durchführung und – nach Möglichkeit – **akademische Kursakkreditierung** zur Erhöhung der Nachhaltigkeitschancen sowie zur Karriereförderung der Teilnehmenden“

„den formalen **Anschluss InWEnts an eine Universität**, so dass mit der Kursteilnahme eine akkreditierte professionelle Qualifikation erworben werden kann“

„Ein Teilnehmer hat die Aufmerksamkeit auf das Interesse gelenkt, ein **Kreditsystem** einzuleiten, wie es **heutzutage bei USAID funktioniert**. Nachdem man einen Kurs erfolgreich erledigt, bekommt man Kreditpunkte, die bei weiteren Ausbildungen gelten“

„Für die Gewinnung der Teilnehmer war entscheidend: (...) Die Zahl von **Fortbildungspunkten**, die Teilnehmer durch die Absolvierung des Kurses erwerben konnten“

„Die **Akkreditierung**, die dem Trainingskurs **durch die UWC** [University of Western Cape] verliehen wurde, hat die Motivation und Wertschätzung des Kurses seitens der Teilnehmer/innen sowie ihre Karriereentwicklung gefördert“

In fast der Hälfte aller Evaluationsberichte wird aufgeführt, dass eine **systematische Nachbetreuung** der Alumni in den evaluierten Programmen *nicht* gegeben war. Nachkontakte sind jedoch von zentraler Bedeutung, um Wirkungen auf institutioneller Ebene zu stärken und die *Nachhaltigkeit* der Wirkungen auf individueller Ebene zu sichern. Von den Evaluatoren/innen werden diverse Nachbetreuungsmöglichkeiten angeführt: von der Implementierung *virtueller Chat-Räume* oder lediglich der Möglichkeit, Fragen per Email an einen *Mentor* stellen zu

können, über *Auffrischkurse*, die neuen Entwicklungen behandeln, oder *Konferenzen und Seminare* bis hin zu individuellem *Coaching* und *on-the-job Unterstützung*. Insbesondere letzteres scheint von Bedeutung: Die klassische Alumni-Arbeit von InWEnt beinhaltet in erster Linie Nachbetreuung (oder besser Nachkontakt) i.S. einer Vernetzung ehemaliger Teilnehmer/innen, **individuelles Coaching und Mentoring** ist dagegen nicht grundsätzlich vorgesehen. In einem der analysierten Berichte wird ein Programm evaluiert, dass ein solches individuelles, intensives Coaching- und Mentoring implementiert hatte und hierdurch entscheidende Nachhaltigkeit bewirken konnte.



Auszug: Nachbetreuung, Follow-up & Coaching Konzepte

„The module had **no direct follow-up activity**. (...) The outcome and possibly also the impact could be maximized if refresher courses were to be conducted in light of (...) new developments. It would also enhance the **sustainability** of the module if an e-platform for discussions and exchange of practical experience among staff members could be facilitated“

„Development of some sort of **post-training assistance**: help-line, pre-scheduled follow-up visits, Q&A online forum (if internet connectivity is available), additional readings“

„the claims for **follow-up** and **on-the-job support action** are striking. Obviously this had not been anticipated by the course organizers, thus no follow-up activities were carried out“

„Um in diesem Feld signifikante Veränderungen zu bewirken, sind **Folgemaßnahmen** in Form von **Coaching** und Begleitung notwendig“

„Entwicklung eines **Follow-up Coaching Konzepts** für von Teilnehmenden definierte Projekte, z.B. durch systematische Aufarbeitung von Beiträgen in Chats, in denen die Teilnehmenden von praktischen Problemen bei der Umsetzung eigener Projekte berichten“

„Es wird empfohlen (...) ein **Coaching Konzept zu entwickeln**“

„Sowohl auf die Individuen als auch vor allem auf der Institutionenebene sollte die Wirkung durch eine **Nachbetreuung** verstärkt werden“

„Ebenfalls deutlich wurde die Notwendigkeit, im Interesse von institutioneller Nachhaltigkeit und Effizienz (..) die Bemühungen um ein **strukturiertes Nachbetreuungsprogramm** mit Unterstützung der Alumni **am Arbeitsplatz (coaching)** zu intensivieren“

„Alle befragten Teilnehmer äußerten den Wunsch nach **stärkerer Nachbetreuung**, kontinuierlicherer Information und gezieltem Kurzzeittraining“

„The InWEnt training model qualitatively differs from the traditional cascade model or ‘quick fix’ training frequently offered in developing countries due to a lack of adequate resources. The added value of the InWEnt training model lies in the three-year training **followed by intense coaching and mentoring** of trainers of trainers (multipliers), facilitators and school principals and teachers. (...) This **post-training-phase mentoring** created a context for the experts and adult learners, trainers of trainers (multipliers) and facilitators, to critically review their own training strengths and weaknesses and find immediate remedies for the problems encountered“



Weitere zentrale **Erfolgsfaktoren** für die Programmarbeit sind...

- ... die **Akkreditierung** der InWEnt-Kurse von Hochschulen der Partnerländer, um die „Wertschätzung“ der Kurse zu erhöhen
- ... **individuelles Coaching & Mentoring**, um die Nachhaltigkeit zu sichern

In einigen der analysierten Evaluationsberichte wurde auf **einzelne Aspekte** hingewiesen, die möglicher Weise ebenfalls **Erfolgsfaktoren** darstellen könnten:

So wurde in drei Berichten das Problem der Fluktuation fortgebildeter Mitarbeiter/innen negativ erwähnt. Ein Bericht dagegen führt aus, dass im Rahmen des Programms bei den Behörden generell um **Verständnis geworben** wurde, ausgebildete Lehrkräfte für eine **Zeitraum von 5 Jahren nicht zu versetzen**, was offensichtlich dazu geführt hatte, dass in diesem Fall die Fluktuationsrate äußerst gering gehalten werden konnte, was wiederum für Wirkungen auf

organisatorischer Ebene wichtige Voraussetzung zu sein scheint. Eine solche *Übereinkunft* könnte im Kontext der oben dargestellten *verstärkten Kooperation und Abstimmung* mit den EZ-Partnern und Stakeholdern vor Ort gut umgesetzt werden.

In einem der analysierten Berichte wurde ein weiterer Punkt positiv hervorgehoben: Hier wurde dezidiert erwähnt, dass das evaluierte (noch nicht nach PriME geplant) Programm beispielhaft *umgesteuert* und *neu ausgerichtet* wurde und dass diese Steuerungsentscheidung entscheidender Erfolgsfaktor für positive Outcomes gewesen sei. Da die Umsteuerung als „mutig“ bezeichnet wird, könnte abgeleitet werden, dass derartig notwendige Schritte nicht immer angegangen werden.

2.5 Wirkungsmodell

Zwar zeigen sich wie oben ausgeführt wurde auf organisatorischer, institutioneller Ebene zum Zeitpunkt der Abschlussevaluierungen eher geringe Wirkungen. In einem Bericht wird allerdings angemerkt, dass die *Wirkungen auf individueller Ebene* weit über die intendierten Ziele hinausgehen – dass sich grundlegende und insbesondere *dauerhafte* Änderungen bei den Teilnehmer/innen eingestellt hätten, die in *jeder* Anstellung bei *jeder* Organisation früher oder später Wirkungen entfalten. Es kann also vermutet werden, dass Wirkungen auf organisatorischer Ebene bzw. organisatorischen Ebenen allgemein einen längeren Zeitraum benötigen, um sich zu entfalten, und daher im Kontext der Abschlussevaluationen schlichtweg (noch) nicht nachweisbar waren.

Grundsätzlich ist zu bedenken, dass Wirkungen im Rahmen von Evaluierungen letztendlich nur auf Basis von *Wirkungsmodellen* angemessen nachgewiesen werden können. Eine allgemeine Wirkungskette zu Capacity Building mit wissenschaftlich fundierten Ursache-Wirkungs-Hypothesen u.a. abgeleitet aus der Bildungsforschung fehlt bislang allerdings noch. Es ist daher zu überlegen, ob nicht InWEnt als international renommierter Anbieter von Capacity Building ein solches *allgemeines Wirkungsmodell zu Capacity Building* als Grundlage für Wirkungsevaluierungen erarbeiten (lassen) sollte.



Ein allgemeines *Wirkungsmodell zu Capacity Building* als Grundlage für Wirkungsevaluierungen ist auf Basis wissenschaftlicher Erkenntnisse zu erarbeiten.