



Universität Potsdam



Jan Lorenz Wilhelm (Hrsg.)

Evaluation komplexer Systeme

Systemische Evaluationsansätze in der deutschen
Entwicklungszusammenarbeit

Potsdamer Geographische Praxis // 10

Potsdamer Geographische Praxis // 10

Jan Lorenz Wilhelm (Hrsg.)

EVALUATION KOMPLEXER SYSTEME

Systemische Evaluationsansätze in der deutschen
Entwicklungszusammenarbeit

Universitätsverlag Potsdam

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de/> abrufbar.

Universitätsverlag Potsdam 2015

<http://verlag.ub.uni-potsdam.de>

Am Neuen Palais 10, 14469 Potsdam
Tel.: +49 (0)331 977 2533 / Fax: -2292
E-Mail: verlag@uni-potsdam.de

Die Schriftenreihe *Potsdamer Geographische Praxis* wird herausgegeben vom Institut für Geographie der Universität Potsdam.

ISSN (print) 2194-1599
ISSN (online) 2194-1602

Das Manuskript ist urheberrechtlich geschützt.
Gestaltung: André Kadanik, Berlin
Satz: Ute Dolezal
Titelfotos: © klaas hartz / pixelio.de
Druck: docupoint GmbH Magdeburg

ISBN 978-3-86956-336-7

Zugleich online veröffentlicht auf dem Publikationsserver der Universität Potsdam:
URN <urn:nbn:de:kobv:517-opus4-78384>
<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-78384>

INHALTSVERZEICHNIS

Jan Lorenz Wilhelm

**EVALUATIONEN KOMPLEXER SYSTEME:
EIN FELD MIT FRAGEZEICHEN..... 7**

- 1 Evaluationen: Kein einfaches Unterfangen..... 9**
- 2 Evaluationsdebatte:
Von Linearität zur Komplexität 12**
- 3 Exkurs:
Systemische Ansätze in Therapie und Beratung..... 19**
- 4 Entdeckung systemischer Ansätze für die Evaluation 24**

Richard Hummelbrunner | Bernward Causemann | Theo Mutter | Michaela Raab

SYSTEMISCHE ANSÄTZE IN DER EVALUATION..... 33

- 1 Einleitung 35**
- 2 Die Hauptbestandteile systemischer Ansätze in Evaluationen 38**
- 3 Evaluation aus systemischer Sicht 57**
- 4 Wirkungen und Kausalität aus systemischer Sicht..... 61**
- 5 Mehrwert und Rahmenbedingungen systemischer Ansätze..... 74**

Anja Bugenhagen

**„OBJEKTIV GEHT NICHT MEHR.“ – EXPERTENINNENSICHT
AUF SYSTEMISCHE EVALUATIONSANSÄTZE IN DER
ENTWICKLUNGSZUSAMMENARBEIT..... 81**

1	Einleitung	83
2	Herausforderungen von Evaluationen in der Entwicklungszusammenarbeit.....	84
3	Zwischenfazit	94
4	Bedeutungsaufladung systemischer Evaluationsansätze und ihre Anwendung in der Praxis	96
5	Chancen und Grenzen systemischer Evaluationsansätze	110
6	Zusammenfassung und Fazit	120

Hendrik Smid

**LOST IN COMPLEXITY – HOW TO EVALUATE SUSTAINABLE
DEVELOPMENT? 127**

1	Background: Sustainable development and evaluation	129
2	Empirical analysis of perspectives on evaluation in German development cooperation.....	134
3	Discussion – What is not systemic?	160
4	Conclusion	164

EVALUATIONEN KOMPLEXER SYSTEME: EIN FELD MIT FRAGEZEICHEN

Jan Lorenz Wilhelm

1 EVALUATIONEN: KEIN EINFACHES UNTERFANGEN

In den letzten Dekaden ist das Instrument der Evaluation zu einem festen Bestandteil der Politikgestaltung und -steuerung avanciert (vgl. STOCKMANN, MEYER 2010: 35). Unter Evaluation wird dabei im Allgemeinen die Bewertung eines bestimmten Sachverhalts nach festgelegten Kriterien verstanden. Dieser „Boom der Evaluation“ (ebd.: 38) lässt sich nicht auf einzelne Politikfelder begrenzen, denn sowohl in der Entwicklungspolitik, der Bildungspolitik, der Hochschulpolitik, der Städtebauförderung und in der Ausrichtung des Verwaltungshandelns – um nur einige Beispiele zu nennen –, haben Evaluationen kontinuierlich an Bedeutung gewonnen (vgl. u. a. BOGUMIL et al. 2004: 9; SEDLACEK 2004; GRANZER et al. 2008; KEMPFFERT, ROLF 2008).

In der wissenschaftlichen Debatte wird die zunehmende Bedeutung von Evaluation häufig mit den fortdauernden Defiziten öffentlicher Haushalte in einen Zusammenhang gestellt (vgl. ROMMELSPACHER 2001: 6; BOGUMIL et al. 2004; STOCKMANN, MEYER 2010). Aufgrund des stärker werdenden Wettbewerbs um finanzielle Mittel, erwächst seitens der Politik und Verwaltung ein Interesse an der Verbesserung und an einer besseren Rechtfertigung des Verwaltungshandelns (vgl. BOGUMIL et al. 2004: 9). Des Weiteren wird die Bedeutungszunahme von Evaluation mit der Veränderung von Leitbildern und Konzepten des New Public Management und der damit einhergehenden Ergebnisorientierung öffentlichen Handelns erklärt (vgl. WOLLMANN 2005: 502 ff.). Begriffe wie Evidence Based Management beschreiben Idealbilder, wonach sich das Handeln der Politik und Verwaltung an seiner Wirksamkeit ausrichten soll. Dies wiederum setzt voraus, dass Informationen über die Wirksamkeit öffentlichen Handelns – also Evaluationsergebnisse – vorliegen (vgl. BOGUMIL et al. 2004: 9; STOCKMANN, MEYER 2010: 51).

Ein Blick auf das Handlungsfeld der deutschen Entwicklungspolitik macht nun im Hinblick auf den Stellenwert von Evaluationen zwei Aspekte deutlich. Zum einen fällt auf, dass Evaluationen in diesem Handlungsfeld eine besonders lange Tradition haben. Während beispielsweise im Gesundheitswesen und in der deutschen Städtebauförderung Evaluationen erst in der zweiten Hälfte der 1990er Jahre systematisch programmkonzeptionell Einzug hielten¹, waren Evaluationen in der Entwicklungspolitik bereits in den 1980er fest etabliert. Zum anderen fällt auf,

¹ In der deutschen Städtebauförderung wurde erst 1998 mit dem Bund-Länder-Programm „Soziale Stadt“ das erste Förderprogramm verabschiedet, in welchem das Instrument der Evaluation von Beginn an konzeptionell verankert wurde (vgl. WILHELM 2012: 39 ff.).

dass Evaluationen in diesem Politikfeld seit der Jahrtausendwende noch einmal an Relevanz gewonnen haben. Diese Bedeutungszunahme resultiert sowohl aus der politischen Agenda, sichtbare und nachweisbare Erfolge zu den verabschiedeten Millenniumsentwicklungszielen erreichen zu wollen, als auch aus der großen wissenschaftlichen und medialen Kritik (siehe BUGENHAGEN in diesem Band). Die Gründung des Deutschen Evaluierungsinstituts der Entwicklungszusammenarbeit (DEval) kann als Indiz dafür genommen werden. *Kurzum*: ein Instrument scheint sich ganz offensichtlich zu bewähren.

Nun stellt es sich aber gar nicht so einfach dar, eine Evaluation durchzuführen, also einen bestimmten Sachverhalt nach festgelegten Kriterien zu bewerten. Und so wurde dieser Evaluationsboom von Anfang an von einer wissenschaftlichen Debatte begleitet, in der die Herausforderungen und die Logik von Evaluationen kritisch diskutiert werden (vgl. WILHELM 2012: 54 ff.; ROLFES, WILHELM 2014: 26). An dieser Stelle werden ohne Anspruch auf Vollständigkeit einige dieser diskutierten Herausforderungen skizziert:

- » *Herausforderung unpräziser Zielformulierungen:*
In Evaluationen geht es häufig darum, die Zielerreichung von Programmen, Maßnahmen und Projekten zu ergründen. Als Schwierigkeit wird nun in der Evaluationsdebatte markiert, dass Zielformulierungen häufig nur unpräzise vorlägen.
- » *Herausforderung unzureichender Datengrundlage:*
Die Güte von Evaluationen hängt von der Verfügbarkeit der zu interpretierenden Daten ab. Wenn ein Förderprogramm zur Verringerung von Arbeitslosigkeit in einem Stadtteil beitragen soll, dann könnte es das Ziel einer Evaluation sein, die Veränderung dieser Arbeitslosigkeit zu bewerten. Häufig liegen die im Rahmen einer Evaluation ‚gewünschten‘ Daten (in diesem Fall Daten zur Arbeitslosenquote) nicht flächenspezifisch (hier für das Wohngebiet) vor und müssen geschätzt oder – falls datenschutzrechtlich erlaubt – aufwendig neuerhoben werden.
- » *Herausforderung der Kausalitätszuschreibungen:*
Auch wenn im Rahmen einer Evaluation die von einem Programm, Maßnahme oder Projekt intendierten Ziele rekonstruiert würden, Datenreihen vorlägen und Aussagen über Veränderungen möglich erschienen – z. B. eine Verringerung der Arbeitslosenquote im Wohngebiet während der Interventionslaufzeit um zwei Prozentpunkte – dann bestünde nach wie vor eine Herausforderung in der Festschreibung klarer Kausalzusammenhänge. Die Frage könnte dann also sein, ob die Verringerung der Arbeitslosenquote als Folge der Interventionen oder in Folge gesamtwirtschaftlicher Entwicklungen gesehen werden kann.

» *Herausforderung strategischer Kalküle:*

Studien zeigen, dass ein polit-strategischer Umgang mit Evaluationen eine beachtenswerte Rolle spielt. Die an Evaluationen Beteiligten wissen in der Regel sehr genau, wie in dem jeweiligen politischen Handlungsfeld – egal ob Stadt- und Regionalentwicklung oder Entwicklungszusammenarbeit – mit Evaluationen umzugehen ist und stellen sich auch auf diese Evaluationserfordernisse ein. Der Tatsache, dass Evaluationen durchgeführt werden müssen und eine mehr oder weniger zentrale Rolle im Hinblick auf Legitimierung, Kontrolle und Monitoring spielen, sind sich alle Akteure bewusst. Sie nutzen, instrumentalisieren und interpretieren die Evaluationen in ihrem Sinne. STOCKMANN und MEYER (2010: 75) haben dies als taktische Evaluationsfunktion beschrieben und WILHELM (2012: 223) spricht daran anknüpfend von der Symbolisierungsfunktion von Evaluationen.

Angesichts dieser vielfältigen Herausforderungen, möchte dieser Beitrag weitere Gestaltungsmöglichkeiten von Evaluationen erkunden. Dazu wird zunächst auf wissenschaftstheoretische Grundpositionen in der Evaluationsdebatte eingegangen und die Relevanz von konstruktivistischen Positionen herausgestellt. Daran anschließend wird ein Blick auf systemische Therapie und Beratung geworfen. Ein Feld, aus welchem sich Methoden, Haltungen und Techniken für Evaluationen ableiten lassen, die allesamt auf der Annahme sozialer Komplexität basieren. Vor diesem Hintergrund werden abschließend die bisherige Auseinandersetzung mit systemischen Ansätzen in der Evaluation deutscher Entwicklungszusammenarbeit skizziert und die weiteren Beiträge des vorliegenden Sammelbandes vorgestellt und thematisch zugeordnet.

2 EVALUATIONSDEBATTE: VON LINEARITÄT ZUR KOMPLEXITÄT

Es besteht Konsens darüber, dass die Umsetzung von Evaluationen mit Herausforderungen und Schwierigkeiten einhergeht. Wie nun aber mit diesen Herausforderungen und Schwierigkeiten umgegangen werden kann, hängt entscheidend davon ab, mit welchem Wirklichkeitsverständnis auf Evaluation geschaut wird. Und genau dies gilt es im Folgenden in den Blick zu nehmen.

ZUNÄCHST: SCHEINBARE LINEARITÄT UND STEUERUNGSOPTIMISMUS

Auf der Suche nach prägenden Wirklichkeitsverständnissen zeigt sich, dass in Evaluationsvorhaben häufig ein positivistisches Wirklichkeitsverständnis zugrunde gelegt wird. Dieses äußert sich auf zweierlei Weise. Zum einen basiert eine Vielzahl von anwendungsbezogenen Vorschlägen zur Umsetzung von Evaluationen auf der Annahme, dass den Evaluationsgegenständen (Programme, Projekte, Maßnahmen) eine Programmtheorie und ein daraus abgeleitetes Handlungskonzept zugrunde gelegt wird (vgl. WOLLMANN 2000b: 202; KROMREY 2001: 109 ff. und 116; STOCKMANN, MEYER 2010: 170). Was meinen diese Begriffe? – Eine Programmtheorie wird dabei als theoretischer Erklärungs- und Lösungsansatz eines Sachverhaltes verstanden. Er beschreibt einen Kurs, der notwendig erscheint, um einen Sachverhalt zu verbessern, Interventionen sinnvoll erscheinen zu lassen. In dieser Programmtheorie wird ausgearbeitet, warum der beschriebene Sachverhalt so ist wie er ist und welche Ziele angestrebt werden sollen, damit Veränderungen eintreten können. Die Ziele werden dann bestenfalls in einem Zielsystem (vgl. Abb. 1) verschiedenen Hierarchieebenen zugeordnet. In der Literatur wird zwischen Leitzielen, Mittlerzielen und Handlungszielen unterschieden, wobei ein Zielsystem durchaus auch mehr als drei Ebenen aufweisen kann (vgl. BEYWL, SCHEPP-WINTER 2000: 10).

„Der Zweck des [...] Zielsystems liegt darin, von allgemeinen, übergeordneten, weitgreifenden Zielen Verbindungen zu spezifischen, konkreten und kurzfristigen Zielen herstellen zu können und sich umgekehrt zu fragen, mit welchen übergeordneten Zielen die konkreten Ziele in Verbindung stehen“ (BEYWL, SCHEPP-WINTER 1999: 42).

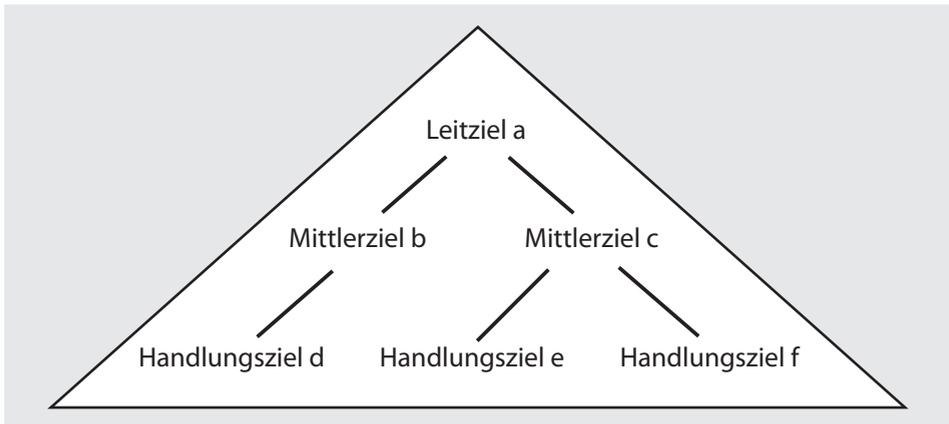


Abb. 1: Die Vorstellung eines linear zusammenhängenden Zielsystems
 Quelle: eigener Entwurf nach BEYWL und SCHEPP-WINTER (2000: 10)

Bildlich gedacht hat das Zielsystem bestehend aus Leit-, Mittler- und Handlungszielen die Form einer Pyramide (vgl. Abb. 1). BEYWL und SCHEPP-WINTER gehen davon aus, dass die Ermittlung einer Zielerreichung davon abhängt, wie konkret ein Ziel formuliert wurde. Die Erreichung von Handlungszielen könne gut und die von Leitzielen nur begrenzt überprüft werden (vgl. BEYWL, SCHEPP-WINTER 2000: 10). Ein Handlungskonzept folge nun der Logik einer Programmtheorie. Dies bedeute, dass die in einem Konzept aufgeführten Programme, Projekte, Maßnahmen gewissermaßen die Operationalisierung der theoretischen Handlungsziele darstelle. Das Konzept beschreibe die Umsetzung des durch die Theorie vorgegebenen Kurses. Unabhängig davon, wie detailliert eine Programmtheorie oder ein daraus abgeleitetes Handlungskonzept sei, könne sich eine Bewertung eines Programms, eines Projektes oder einer Maßnahme immer auf die scheinbar zugrunde gelegte Programmtheorie und das Handlungskonzept beziehen. Zur Aufgabe der Evaluation werde es dann, die zugrunde gelegten Annahmen linearer Wirkungszusammenhänge anhand von indikatorbasierten Zielüberprüfungen nachzuweisen.

Zur Verdeutlichung wollen wir uns an dieser Stelle ein konkretes Beispiel denken (vgl. Abb. 2). Wir stellen uns einmal vor, dass eine öffentliche Verwaltung das Umfeld eines Wohngebietes (Sachverhalt a) anhand einer Spielplatzneugestaltung (Intervention b) und einer aufwändigen Pflanzaktion (Intervention c) aufwerten wollte. Eine durchzuführende Evaluation könnte sich nun zur Aufgabe machen, auf der Basis der zugrunde gelegten Programmtheorie und der darin enthaltenen Wenn-Dann-Beziehungen nachvollziehbar aufzuzeigen, welche Auswirkungen die Intervention b (Spielplatzneugestaltung) und die Intervention c (Pflanzaktion) für den Sachverhalt a (Wohnumfeld) hatten.

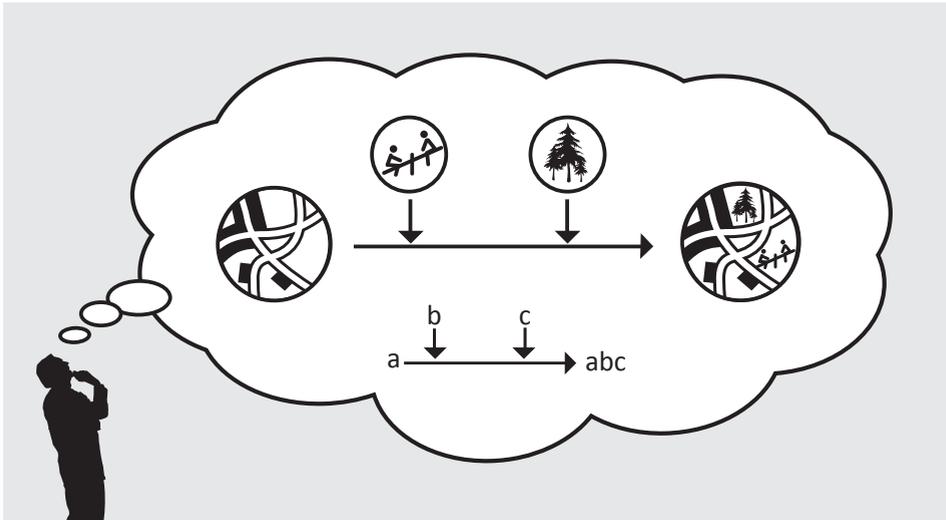


Abb. 2: Evaluation als Instrument zur Darstellung linearer Wirkungszusammenhänge; Quelle: WILHELM 2012: 21

Zusammenfassend entsteht zunächst der Eindruck, dass sich während der Umsetzung von Evaluationen Annahmen linearer Wirkungszusammenhänge konzeptionell niederschlagen. Dieses positivistische Wirklichkeitsverständnis zeigt sich aber nicht nur in anwendungsbezogenen Vorschlägen zur Umsetzung von Evaluationen, sondern auch in Abhandlungen, die den Sinn und Zweck von Evaluationen beschreiben. So heißt es in einer Stellungnahme der Bundesregierung Deutschland, dass „Evaluationen [...] über Effektivität, die (beabsichtigten und unbeabsichtigten) Wirkungen sowie die Wirtschaftlichkeit von Maßnahmen oder Programmen [informieren]“ (BUNDESREGIERUNG 2006: Kap. 1). Durch Evaluationen erzielte Ergebnisse, Schlussfolgerungen und Empfehlungen – so heißt es weiter – müssen nachvollziehbar sein und auf validen Daten beruhen. Evaluationen dienen dazu, Prozesse transparent zu machen, Wirkungen zu dokumentieren und Zusammenhänge aufzuzeigen und auch alternative Handlungsoptionen zur Erreichung derselben Ergebnisse zu beleuchten (vgl. ebd.).

STOCKMANN bestätigt dieses positivistische Wirklichkeitsverständnis, indem er die Einschätzung anführt, dass die „empirischen Methoden zur Informationsgewinnung und systematische Verfahren zur Informationsbewertung anhand offen gelegter Kriterien [...], [...] eine intersubjektive Nachprüfbarkeit möglich machen“ (STOCKMANN 2006: 65; vgl. zum positivistischen Wirklichkeitsverständnis auch Tab. 1).

UND DANN: OFFENSICHTLICHE KOMPLEXITÄT UND STEUERUNGSPESSIMISMUS

In der Evaluationsforschung wird das in Evaluationen zugrunde gelegte Wirklichkeitsverständnis jedoch vielfältig diskutiert. In einer Abhandlung über verschiedene methodische Paradigmen in der Evaluationsforschung kommt STOCKMANN zu folgender Aussage:

„Der methodologische Rigorismus wurde zudem vom konstruktivistischen Paradigma in Frage gestellt. Das Vorhandensein einer einzigen ‚wahren‘ Realität wird bestritten. Stattdessen wird angenommen, daß Realität aus verschiedenen Perspektiven sozial konstruiert ist, die in Konflikt zueinander stehen können. [...] Nach der erkenntnistheoretischen Position des Konstruktivismus sind grundsätzlich keine Aussagen über die tatsächliche Beschaffenheit der Welt möglich [...]“ (STOCKMANN 2004: 21).

Nach dem konstruktivistischen Paradigma kann es der Zweck von Evaluation nicht sein, eine möglichst ‚realitätsnahe‘ Darstellung und Bewertung der vorgefundenen Verhältnisse zu liefern (STOCKMANN 2006: 73). Und auch der Evaluationsforscher SEDLACEK verweist auf die sich wandelnden Paradigmen:

„Vor dem Hintergrund der wissenschaftstheoretischen Diskurse der vergangenen Jahrzehnte ist [...] von der Vorstellung einer wissenschaftlichen ‚Objektivität‘ ohnehin Abschied zu nehmen. [...] Der Sinn von Evaluation liegt damit – diesbezüglich scheint mir ein Konsens vorhanden – im gemeinsamen Lernen, im gemeinsamen Erarbeiten einer neuen Sicht der Dinge, und das heißt in der Umdeutung, dem Reframing bisheriger Wirklichkeitsauffassungen“ (SEDLACEK 2004: 23 f.).

An dieser Stelle ließen sich weitere Beispiele für eine zaghafte Kritik der linearen Paradigmen finden. Zaghafte deshalb, weil zum einen herausgestellt wird, dass die in Evaluationen häufig zugrunde gelegten linearen Kausalbeziehungen der beobachteten Komplexität von Organisationen nicht gerecht werden, zum anderen jedoch an diese Feststellung die Forderung nach ganzheitlicheren Wirkungsanalysen und der Schaffung einer objektiveren Basis für zukünftige Evaluationen anschließen und somit eine eher positivistische Perspektive zum Ausdruck kommt.

Konsequent-konstruktivistische Auseinandersetzungen mit dem Thema Evaluation stellen eher die Ausnahme dar. Drei Beispiele aus der deutschsprachigen Evaluationsdebatte liefern BLANK (2008), BAECKER (2008) und WILHELM (2012). Sie ordnen sich ein in eine Evaluationsforschung, die sich dem Phänomen Evaluation aus einer systemtheoretischen Perspektive – also einer radikal konstruktivistischen Perspektive – nähern. BLANK (2008) leitet in Rückgriff auf die Systemtheorie

Luhmannscher Prägung her, dass es sich bei einer Evaluation um eine spezifische Operationsweise eines sozialen Systems handelt. Als Beobachtung von Beobachtung trägt eine Evaluation zur Anpassung systemimmanenter Erwartungsstrukturen – was gleichbedeutend mit Lernen sei – bei (vgl. BLANK 2008: 77 u. 82). Detailreich leitet BLANK her, wie Evaluation in einem systemtheoretischen Verständnis organisationales Lernen begünstigt. Er zeigt, dass den „Erkenntnismöglichkeiten durch Evaluationen prinzipielle epistemologische Grenzen gesetzt“ (BLANK 2008: 95) sind:

„Evaluation kann nur innerhalb der eigenen systemischen Struktur und auf Basis der zum Zweck der Beobachtung gewählten Ausgangsunterscheidung Information erzeugen. Gegenüber anderen Ereignissen bleibt sie blind“ (BLANK 2008: 95).

In diesem Zusammenhang empfiehlt Blank, „sich von der Idee zu verabschieden, mittels Evaluation sei eine objektivierbare Annäherung an den beobachteten Gegenstand prinzipiell möglich“ (BLANK 2008: 95). BAECKER bespricht in seinem Beitrag am Beispiel der Evaluation kultureller Projekte eine konstruktivistische Methodologie (vgl. BAECKER 2008: 99 f.). Basierend auf den formtheoretischen Überlegungen von SPENCER BROWN (1977) entwickelt BAECKER ein Formmodell von Evaluation, welches systematisch die vielfältigen Evaluationskontexte einbezieht. Charakteristisch für das Formmodell ist, dass es ganz bewusst einen blinden Fleck mitführt. Evaluation erscheint dadurch als nicht eindeutig lösbar sondern stets so-oder-auch-anders darstellbar. BAECKER versteht in diesem Sinne Evaluation „als ein Gesprächsangebot zwischen Auftraggeber beziehungsweise Durchführer der Evaluation auf der einen Seite und dem evaluierten Projektträger auf der anderen Seite“ (BAECKER 2008: 97). WILHELM schließt an die Überlegungen von BAECKER und BLANK an und leitet aus systemtheoretischer Perspektive ab, welche Funktionen Evaluationen für Organisationssysteme erfüllen (vgl. WILHELM 2012). WILHELM zur Folge können Evaluationen in Organisationen dazu dienen, Komplexitätsreduktionsschemata zu aktualisieren, politische Steuerungsintentionen zu visualisieren und als sinnvoll erscheinen zu lassen sowie Lernen und Reflexivität zu symbolisieren (vgl. WILHELM 2012: 195 ff.)

Diese wissenschaftliche Positionen zeichnet zusammenfassend aus, dass das Handeln eines sozialen Systems als Black Box betrachtet werden kann, zu deren Funktionsweise keine klaren Wenn-Dann-Aussagen, also keine eindeutigen Aussagen im Sinne $a + b + c = abc$ (vgl. Abb. 2), mehr zulässig erscheinen (vgl. v. AMELN 2004: 85 ff.). EGNER und RATTER fassen demzufolge zusammen, dass es sich bei sozialen Systemen „um nicht-lineare, dynamische Systeme handelt, die fern von einem Streben nach Gleichgewicht sind und deren Verhalten durch Emergenz und der Systemablauf durch Sprünge und Überraschungen geprägt ist“ (EGNER, RATTER 2008: 11).

Unser Beispiel aus der Wohnumfeldaufwertung aufgreifend, bleibt zunächst sehr fraglich, aus welcher Perspektive und wozu das Wohnumfeld und die durchgeführten Maßnahmen in einen Zusammenhang gestellt werden können (vgl. Abb. 3).

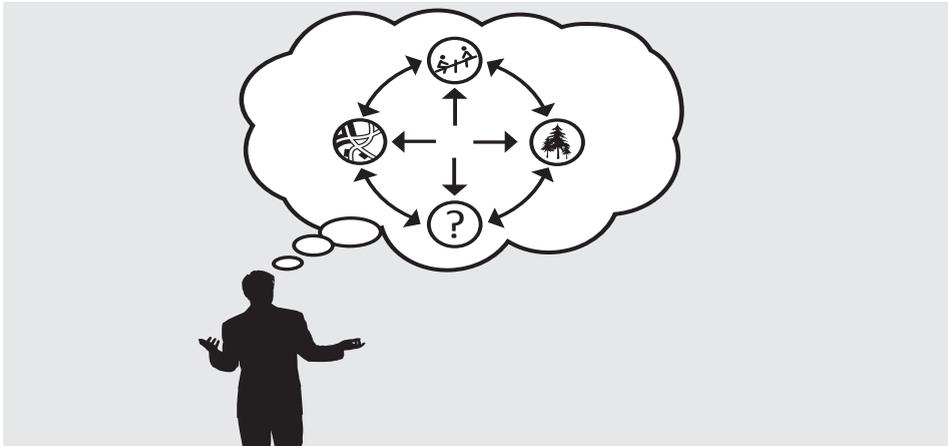


Abb. 3: Evaluationen stellen sich als multikomplex dar
Quelle: eigene Darstellung in Anlehnung an KÖNIGSWIESER, LANG (2008: 39)

UND WAS FOLGT NUN?

In der Evaluationsdebatte schlugen sich diese planungs- bzw. steuerungsskeptischen Annahmen u. a. in der von GUBA und LINCOLN (1989) ins Leben gerufenen Bezeichnung einer *fourth generation evaluation* oder der von FETTERMANN (1994) mitgeprägten *empowerment evaluation* nieder. Während die Evaluationen der ersten drei Generationen auf einem positivistischen Wissenschaftsverständnis beruhen, distanzieren sich Evaluationen der vierten Generation von diesen Vorstellungen und knüpfen an konstruktivistische, hermeneutische und interpretative Ansätze an (vgl. LAY, PAPADOPOULOS 2007: 495). Diese Evaluationsansätze beziehen sich folglich auf verschiedene Denkansätze seit den 1970er Jahren, die mit den Grundannahmen positivistischen Wirklichkeitsverständnisses brechen, deren Anspruch auf Objektivität verneinen und von der eingeschränkten Messbarkeit sozialer Wirklichkeit ausgehen (vgl. STOCKMANN, MEYER 2010: 109 f. und 139). Wissenschaftliche Arbeiten zur vierten Generation erschienen bereits in den 1980er und 1990er Jahren und wurden von GUBA und LINCOLN (1989), PATTON (1987), STAKE (1983) und HOUSE (1993) geprägt (vgl. MERTENS 2004: 48).

Eine konstruktivistische Grundauffassung – so viel ist klar – erfordert „bei Evaluationen eine andere wissenschaftliche Vorgehensweise als im Rahmen einer analytisch-nomologisch orientierten Erfahrungswissenschaft“ (STOCKMANN 2004: 21). Auf welche Art und Weise sich die wissenschaftliche Vorgehensweise auf der Basis eines konstruktivistischen Paradigmas verändern könnte, bleibt eine spannende

Forschungsfrage. Manchmal entsteht der Eindruck, als beschränke sich die Konsequenz auf die Verwendung qualitativer Methoden empirischer Sozialforschung. So führt STOCKMANN bspw. aus, dass „die Anhänger [des konstruktivistischen Paradigmas: der Verf.] ein ‚qualitatives‘ Denken [fordern], um die verschiedenen Sichtweisen und Interpretationen der ‚Realität‘ erfassen zu können“ (STOCKMANN 2004: 21). Und auch Mertens führt an, dass das konstruktivistische Paradigma insbesondere zur Berücksichtigung qualitativer Methoden führte (vgl. MERTENS 2004: 48). Als Beispiele für die angedeuteten Methoden ließen sich vermutlich das FIWE-Schema² und die von Peel skizzierte Sensibilisierung von Evaluationsperspektiven³ anführen (vgl. PEEL 2003: 159 f.).

Für die Suche nach weiteren, auf konstruktivistischen Paradigmen basierenden Methoden und Ansätzen für die Evaluation, erweist sich eine Auseinandersetzung mit dem im deutschsprachigen Raum seit den 1980er Jahren sich formierenden Feld der systemischen Therapie- und Beratungsansätze ertragreich. Dazu soll zunächst das Feld der systemischen Beratung und systemischen Therapie schemenhaft beschrieben werden.

2 FIWE-Schema: Basierend auf der Annahme, dass Bewertungen nicht jenseits von subjektiven Interessen durchgeführt werden können, sollte bei der Durchführung von Evaluationen angestrebt werden, die Evaluierung anhand von Informationen, Daten und Indikatoren sowie insbesondere die Wirkungsmessung für alle Beteiligten so nachvollziehbar wie nur möglich zu gestalten (vgl. BEYWL, SCHEPP-WINTER 1999: 72). Ein systematisches Vorgehen sei hierfür unerlässlich und könnte z. B. mit dem so genannten FIWE-Schema (Fakten – Interpretation – Werte – Empfehlungen) angegangen werden. Zwischen einer daten- und faktenbasierten Beschreibung von Evaluationsergebnissen, der Interpretation dieser Daten und Fakten, der Bewertung dieser beiden Schritte im Hinblick auf die Programmziele und schließlich der Formulierung von Schlussfolgerungen und Empfehlungen sollte klar getrennt werden (vgl. ebd.).

3 Peel schlägt vor, Evaluierungsergebnisse im Lichte der unterschiedlichen Perspektiven, Erwartungen und Motive der Beteiligten zu betrachten und zu interpretieren (vgl. PEEL 2003: 159 f.). Um ein möglichst einheitliches Verständnis über Evaluation und Wirkungsmessung zu erzielen, schlägt Peel vor, die Kooperation und Kommunikation zwischen den Akteuren ausdrücklich zu fordern und hervorzuheben. Dabei seien zum einen die unterschiedlichen Ziele, Begriffsverständnisse, Vorstellungen und Erwartungshaltungen transparent zu machen. Gleichzeitig müsse eine Sensibilisierung der Beteiligten im Hinblick auf ihre eigene Evaluationsperspektive erfolgen (vgl. ebd.).

3 EXKURS: SYSTEMISCHE ANSÄTZE IN THERAPIE UND BERATUNG

War die Bezeichnung ‚systemischer Therapeut‘ oder ‚systemische Beratung‘ noch vor 20 Jahren eher wenigen Personen ein Begriff, so scheint sie seitdem unaufhaltsam auf dem Vormarsch. Bereits 2005 erreichte die Zahl der in Deutschland jährlich in systemischen Ansätzen weitergebildeten Personen die 1000er Marke (vgl. KRIZANITS 2005: 49). Ein Blick in die Ausbildungslandschaft dokumentiert, dass sich diese Dynamik in den letzten zehn Jahren noch einmal beschleunigt hat. Bezeichnend ist dabei, dass dieser Ansatz nicht nur in einer einzelnen Branche auf Interesse stößt, sondern sich vielmehr in ganz unterschiedlichen Bereichen wie Familientherapie, Organisationsberatung, Managementberatung, Politikberatung, in der Sozialen Arbeit, in der Schule, in der Kommunal- und Regionalentwicklung und eben auch zur Durchführung von Evaluationen verbreitet (vgl. auch v. SCHLIPPE, SCHWEITZER 2007: 245 ff.). Aber was bedeutet systemischer Ansatz und wie ist er entstanden?

Zunächst ist wichtig zu verstehen, dass der systemische Ansatz das Ergebnis eines lebhaften Diskussionsprozesses innerhalb der Branche der TherapeutInnen und BeraterInnen ist. Dieser Diskussionsprozess hält nach wie vor an. Ausgelöst wurden diese bis heute aktuellen Diskussionen durch eine Vielzahl von in der zweiten Hälfte des letzten Jahrhunderts ‚populär‘ gewordener wissenschaftlicher Publikationen, die sich allesamt dem systemtheoretischen/konstruktivistischen Theoriegebäude zuordnen lassen. Die an diesem Diskurs beteiligten Wissenschaftler – wie Maturana, Luhmann, von Foerster, Watzlawick – gehörten zwar unterschiedlichen wissenschaftlichen Disziplinen an, standen jedoch in wechselseitigem Kontakt. Sie inspirierten sich gegenseitig und entwarfen somit gemeinsam einen neuen Blick auf Gesellschaft, Organisationen und Personen. Ohne Anspruch auf Vollständigkeit, werden in folgendem Kästchen zentrale wissenschaftliche Positionen dieser Debatte kurz vorgestellt.

WICHTIGE WISSENSCHAFTLICHE POSITIONEN DER SYSTEMTHEORETISCHEN UND KONSTRUKTIVISTISCHEN DEBATTE:

- » Zunächst sei die von verschiedenen Wissenschaftlern ausgearbeitete Idee erwähnt, dass Erkenntnis nur als Folge von Unterscheidungen gedacht werden kann. Der Mathematiker **George Spencer Brown** arbeitete in seinem Werk „Laws of Form“ heraus, dass jede Unterscheidung einen blinden Fleck mitführt, der mit einer weiteren Unterscheidung erhellt werden kann (vgl. SPENCER BROWN 1977). Der Anthropologe, Biologe, Sozialwissenschaftler, Kybernetiker und Philosoph **Gregory Bateson** arbeitete aus, dass am Anfang eines jeden Erkenntnisprozesses Unterscheidungen stehen und dass eine Information nichts anderes sei, als „ein Unterschied, der einen Unterschied ausmacht“ (BATESON 1985: 582). Um erkennen zu können, müssen wir also Stadtteil von Stadt, Verwaltung von Politik und Wirtschaft von Bildung unterscheiden.
- » Der Psychologe **George A. Kelly** entwarf in seinem Werk „Die Psychologie der persönlichen Kontakte“ einen Erklärungsansatz, der beschrieb, dass Menschen entlang von Unterscheidungen versuchen, in ihrem Umfeld konstante Muster zu identifizieren um ihr eigenes Handeln daran auszurichten. Diese Muster – wie Verhaltensmuster im Umgang mit Migranten oder auch Temperaturmuster – sind nicht a priori gegeben, sondern werden von Kelly als Konstrukte bezeichnet (vgl. v. AMELN 2004: 40 ff.).
- » Der Kommunikationstheoretiker **Paul Watzlawick** trug durch seine Arbeiten dazu bei, dass das Sender-Empfänger-Modell relativiert wurde. Wissenspakete – so seine Argumentation – können nicht einfach vom Sender zum Empfänger 1:1 übertragen werden. Auf welche Art und Weise eine Mitteilung – gedacht als Information im Batesons Sinne – verstanden werden kann, hängt stets von den Fähigkeiten und dem Vorwissen des Empfängers ab, muss also als Konstruktionsleistung verstanden werden (vgl. WATZLAWICK et al. 2011).
- » Das Autopoiesis-Konzept des Biologen **Humberto R. Maturana** stellt heraus, dass Systeme operativ geschlossen gedacht werden können. Damit gemeint ist, dass Systeme autonom Informationen – also Reize von außen – verarbeiten und sich dann aus sich selbst heraus weiterentwickeln (Autopoiesis = altgriechisch: autos ‚selbst‘ und poiein ‚schaffen‘).

- » Der Kybernetiker **Heinz von Foerster** erarbeitete in seiner Kybernetik zweiter Ordnung die Idee der reflexiven Selbstreferenz indem er die Wechselwirkung zwischen Steuerndem und Gesteuertem beschreibt: „Sobald [...] jemand ein System beobachtet, es zu verstehen und durch Interventionen steuernd in es einzugreifen versucht, ist er selbst ein Teil des Systemprozesses“ (KÖNIGSWIESER, HILLEBRAND 2005: 27). Heinz von Foerster prägte darüber hinaus den Begriff der Black Box oder der nichttrivialen Maschine. Damit gemeint ist, dass als Folge von Beobachtung und Selbstreflexivität lebende und soziale Systeme undurchschaubar und unvorhersehbar agieren.

- » Der Soziologe **Niklas Luhmann** schließlich erarbeitete in seinem Werk „Soziale Systeme“ eine Theorie, die von einer Ausdifferenzierung der Gesellschaft in Systeme ausgeht, die zueinander ‚autistisch‘ sind und ihre ‚eigenen Logiken‘ verfolgen. Dies meint, dass die jeweiligen Systeme zwar Vorgänge oder Probleme anderer Systeme beobachten können, diese aber nur unter Maßgabe der eigenen Operationslogik und der eigenen Rationalität übersetzen und entsprechend ihrer eigenen Systemzwecke verwenden können. In diesem Sinne schlägt etwa VOGD (2005: 24) vor, grundsätzlich von der Wirksamkeit multipler Logiken auszugehen, diese multiperspektiv zu betrachten und beobachtbar zu machen: Ein Flächennutzungsplan wird bspw. in einer Stadtverwaltung eher im Kontext politischer Planung, in einem Planungsbüro eher im Kontext wirtschaftlicher Kalküle und in einer Universität – sofern er als Veranschauungsmaterial Verwendung findet – im Kontext von Bildung in Erscheinung treten. Soziale Systeme als ein besonderer Typ geschlossener Systeme sind für Luhmann vor allem durch zwei Komponenten gekennzeichnet: Sinnhaftigkeit und Kommunikationsfähigkeit (vgl. LUHMANN 1994: 67). Damit ist ein zentraler Perspektivwechsel verbunden: Soziales wird nicht mehr durch Menschen und Individuen oder durch eine Wechselwirkung unter Individuen erklärt, sondern durch Kommunikation und nur durch Kommunikation. Die luhmannsche Systemtheorie ist also als Gesellschaftstheorie zu verstehen, die die übliche sozial- oder kulturwissenschaftliche Orientierung am Menschen durch eine Orientierung auf Kommunikation bzw. Kommunikationszusammenhänge umstellt.

Besonders löste die konzeptionelle Ausarbeitung des Beobachters oder der Beobachtung und die daran anschließenden Modelle eine intellektuelle Faszination aus. Denn im Begriff der Beobachtung wird sowohl Einsicht (ich erkenne etwas) als auch Blindheit (ich sehe nicht alles), also Möglichkeiten und Limitationen,

gebunden. Und diese Einsicht bzw. Blindheit ist wiederum nur durch weitere Beobachtung sichtbar oder auch nicht sichtbar. Im Beobachterbegriff wird daher ein Maximalmaß an Reflexivität gesehen was insbesondere die Therapie und Beratung vor ganz neue Herausforderungen und auch Chancen stellte.

Aus diesen wissenschaftlichen Positionen ließ sich folglich die Erkenntnis ableiten – Systemiker sprechen dabei gerne auch von einem systemischen Weltbild –, dass es *nicht die eine Wahrheit* gibt. Vielmehr sollte von einer *Vielzahl von Wahrheiten* ausgegangen werden. Diese hängen jeweils von den verschiedenen Sichtweisen, Blickwinkeln, Kontexten ab, in denen sie hergestellt werden und gelten. Alles ist *so oder auch anders* möglich, alles ist komplex. Nichts sollte auf lineare Zusammenhänge reduziert werden, sondern vielmehr kann von Zirkularität und Komplexität ausgegangen werden. Systemiker grenzen gerne ihr systemisches Weltbild von einem positivistischen, auch als mechanistisch bezeichnetem Weltbild ab und stellen dies tabellarisch dar (vgl. Tab. 1).

Positivistisches Weltbild	Systemisches Weltbild
Objektivität, eine Wahrheit, unveränderliche Gesetze	Wirklichkeitskonstruktion, viele ‚Wahrheiten‘, Thesen
Richtig – falsch, schuldig – unschuldig	Kontextabhängigkeit, Nützlichkeit, Anschlussfähigkeit
(Fremd-)Steuerung	Selbststeuerung, Selbstorganisation
Lineare Kausalketten	Vielfältige Wechselwirkungen, Feedbackschleifen
Messbarer, fixer Unterschied	Sich unterscheiden, verändern
Linearer Fortschritt, ändern	Entwicklung, ändern und bewahren, deblockieren
Formale Logik, Widerspruchsfreiheit, Ausschluss	Integration von Widersprüchen, Einbeziehung
Harte Fakten, rationale Beziehungen	Integration von harten und weichen Faktoren (Emotionen, Intuitionen, Kommunikationsprozesse)

Tab. 1: Gegenüberstellung eines positivistischen und systemischen Weltbildes
Quelle: KÖNIGSWIESER, HILLEBRAND 2005: 28 (leicht verändert durch den Verfasser)

Aus diesen neuen Erkenntnissen zur Funktionsweise von Gesellschaft und von Veränderungsprozessen resultierten demzufolge in einigen Zirkeln innerhalb der Branche der Therapeuten und Berater lebhaft Diskussionen. Welche zwingenden Anpassungen für Therapie- und Beratungsansätze müssten entwickelt werden, wenn es kein objektives Richtig und Falsch mehr gab, Wirkungszusammenhänge als komplex zu verstehen waren, keine klaren Ursache-Wirkungs-Beziehungen mehr unterstellt werden können, System-Eigenlogiken bestehen und Kommunikation als Konstruktionsleistung zu verstehen war? Oder anders formuliert:

Welches Potenzial und welche Inspiration ermöglichten die Erkenntnisse über Selbstreferenz, Konstruktion und Perspektivität für die Weiterentwicklung von Beratungsansätzen?

Dabei zeigten sich diese Diskussionen zunächst im nordamerikanischen Raum als besonders produktiv. Als Entwicklungszentren fungierten dabei unter anderem das 1959 gegründete Mental Research Institute in Palo Alto – dort arbeitete auch Watzlawick –, die von Erickson 1957 gegründete Amerikanische Gesellschaft für klinische Hypnose und das von De Shazer 1978 gegründete Brief Therapy Center in Milwaukee. In Nordamerika etablierte sich daraufhin der Begriff Systems Thinking. In den 1980er Jahren entwickelten sich dann im deutschsprachigen Raum zunächst zwei Diskussionscluster mit unterschiedlicher Schwerpunktsetzung: eine familientherapeutische Gruppierung in Heidelberg und eine Gruppierung aus dem Bereich der Gruppendynamik und der Organisationsberatung in Wien und Klagenfurt (vgl. KRIZANITS 2005: 43 f.). In beiden Clustern wurde ein reger Austausch mit den oben erwähnten WissenschaftlerInnen und Therapiezentren sowie anderen Denkschulen – wie die Mailänder Schule zur Familientherapie – gepflegt. Mitte der 1980er Jahre befruchteten dann Vertreter dieser Gruppierungen (z. B. Fritz B. Simon und Gunther Schmidt aus der Heidelberger Gruppe) wiederum andere Netzwerke und begannen Weiterbildungen anzubieten. Aufgrund der zu diesem Zeitpunkt noch überschaubaren Anzahl an beteiligten Personen, konnten sich die Organisationsberater und Therapeuten relativ einfach auf ein „gemeinsames Beratungsparadigma“ (KRIZANITS 2005: 45) festlegen. Und ab Mitte der 1980er Jahre setzte sich dann im deutschsprachigen Raum für diesen Ansatz der Name ‚systemische Beratung‘ durch. Dieser Begriff sei „ein Versuch, die beiden Professionsströme, den klassischen gruppendynamischen-OE-orientierten Ansatz und das, was die systemische Familientherapie an Repertoire zu diesem Zeitpunkt anbieten konnte, zu verbinden“ (WIMMER zitiert nach KRIZANITS 2005: 45 f.). Ergänzend wird aber nach wie vor auch von ‚systemischer Therapie‘ oder allgemeiner von ‚systemischer Intervention‘ gesprochen.

Mittlerweile gibt es eine große Anzahl von Publikationen zum Thema ‚Systemische Ansätze‘ (vgl. z. B. KÖNIGSWIESER, HILLEBRAND 2005; V. SCHLIPPE, SCHWEITZER 2007; KRIZANITS 2009; VARGA V. KIBET, SPARRER 2011). Darin wird jeweils ausgearbeitet, was ein systemisches Wirklichkeitsverständnis bedeutet, welche Grundhaltungen Systemiker verfolgen und welche Methoden und Instrumente zur Gestaltung von Veränderungsprozessen zur Verfügung stehen. Stellt man diese Publikationen nun vergleichend gegenüber, so wird deutlich, dass es nach wie vor unter dem gemeinsamen Dach der systemischen Ansätze eine Vielzahl unterschiedlicher Denkschulen und damit einhergehend unterschiedliche Vorstellungen über die Gestaltung von systemischen Interventionen gibt.

4 ENTDECKUNG SYSTEMISCHER ANSÄTZE FÜR DIE EVALUATION

Wenn es nun also darum geht, auf dem Verständnis sozialer Komplexität aufbauende Methoden und Ansätzen für die Evaluation zu erschließen, dann stellt das Feld der systemischen Ansätze zweifelsohne einen inspirierenden Fundus dar (vgl. WILHELM 2012: 268 f.). Und in der Tat gibt es bereits verschiedene Publikationen, die Verknüpfungsmöglichkeiten zwischen systemischen Ansätzen und der Evaluationsthematik herausstellen (vgl. GALLA et al. 2008; PATTON 2011; WILLIAMS, IMAM 2006; HUMMELBRUNNER 2011).

Bei der Ergründung der Grenzen und Möglichkeiten systemischer Ansätze für Evaluationen geraten schnell zwei Aspekte in den Mittelpunkt:

- » Zum einen stellt sich schnell die Frage nach Ordnung und Übersicht. Das Feld der systemischen Ansätze erscheint sehr breit, vielfältig und auch heterogen. Was berechtigt, von einem Feld und nicht von zwei, drei oder mehr Feldern zu sprechen? Wie wird der Systembegriff ausgelegt, welche Prämissen zugrunde gelegt und welche methodischen Vorgehensweisen daraus abgeleitet?
- » Zum anderen drängt sich die Frage nach dem Unterschied auf. In der Evaluationspraxis gibt es bereits Methoden und Ansätze, die konstruktivistische Positionen aufgreifen (vgl. weiter oben die Debatte über die ‚fourth generation of evaluation‘). Der Ansatz der partizipativen Evaluation bspw. leitet aus dem Aspekt der Perspektivität von Beobachtungen die Evaluationsaufgabe ab, „möglichst viele Beteiligte und Betroffene sowie ihre Perspektiven in einen kommunikativen Aushandlungsprozess zu bringen. In diesem Prozess, dessen Moderator er ist, werden die Ausgangsfragestellungen, das Vorgehen und schließlich die Interpretationen der erhobenen Daten abgestimmt“ (ULRICH, WENZEL 2003: 38 f.). Was wäre an systemischen Ansätzen in der Evaluation folglich neu? Oder führt ein Nachdenken über systemische Ansätze im Rahmen der Evaluationsdebatte nur zur Neubetitelung bereits bewährter Ansätze?

Im deutschsprachigen Raum fand in den letzten Jahren eine intensive Auseinandersetzung mit diesen Fragen statt. Dabei wurden die Möglichkeiten und Grenzen systemischer Ansätze in der Evaluation vorrangig für das Anwendungsfeld ‚Evaluation von deutscher Entwicklungszusammenarbeit‘ ergründet (vgl. folgender Kasten).

ETAPPEN DER AUSEINANDERSETZUNG MIT SYSTEMISCHEN ANSÄTZEN IN DER EVALUATION DEUTSCHER ENTWICKLUNGSZUSAMMENARBEIT

- » **2008: Gründung einer AG „Systemische Ansätze in der Evaluation“**
Innerhalb des AK Entwicklungspolitik der Gesellschaft für Evaluation e.V. wird eine AG gegründet. Diese trifft sich zu 19 Sitzungen und entwirft 2013 ein Arbeitspapier.
- » **2011: Internationale Fachkonferenz: Systemic Approaches in Evaluation**
Am 25. und 26. Januar findet im Rahmen einer internationalen Fachkonferenz der GIZ ein Forum zur Diskussion systemischer Ansätze in der Evaluation statt.
- » **2011: Frühjahrstagung des AK Entwicklungspolitik**
Auf der Frühjahrstagung wird die Fachöffentlichkeit zur vertieften Diskussion systemischer Ansätze eingeladen.
- » **2012: DEval Expertendiskussion**
Das Deutsche Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval) veranstaltet mit Bob Williams und Richard Hummelbrunner eine Veranstaltung zum Thema „Dealing with Complexity“.
- » **2013: Frühjahrstagung des AK Entwicklungspolitik**
Auf der Frühjahrstagung wird das Arbeitspapier der AG „Systemische Ansätze in der Evaluation“ vorgestellt und diskutiert.
- » **2013: Jahrestagung der DeGEval**
Unter dem Leitthema ‚Komplexität‘ werden Themenbereiche wie Mehrperspektivität, Zielkonflikte, Planungsunsicherheit und Entwicklungsdynamiken kontrovers diskutiert und auch Möglichkeiten systemischer Ansätze zur Diskussion gestellt.

Der vorliegende Band möchte nun mit drei ausgewählten Beiträgen Einblicke in diese Debatte über systemische Ansätze in der Evaluation deutscher Entwicklungszusammenarbeit bereitstellen.

In dem Beitrag *Systemische Ansätze in der Evaluation* stellen die Autoren HUMMELBRUNNER, CAUSEMANN, MUTTER und RAAB für die Zielgruppe der EvaluatorInnen und Auftraggeber aus dem Bereich der Entwicklungszusammenarbeit wesentliche Merkmale systemischer Ansätze und Unterschiede zu anderen Evaluationsansätzen heraus. Zur Erläuterung der zentralen Merkmale einer systemischen Herangehensweise werden dabei drei systemische Konzepte (Beziehungen, Perspektiven und Systemgrenzen) in den Mittelpunkt gestellt und exemplarisch mit Methoden verknüpft. HUMMELBRUNNER et al. verweisen auf die Bedeutung einer systemischen Grundhaltung und stellen vier Eckpunkte eines systemischen Evaluationsverständnisses heraus. Nach dieser allgemeinen Annäherung an das Phänomen systemische Ansätze und Evaluation widmen sich die Autoren dem Thema Wirkungen und Kausalität. Sie stellen heraus, wie Wirkungen aus systemischer Perspektive typisiert und welche systemischen Wirkungsanalysen und Kausalanalysen identifiziert werden können. Abschließend betrachten die Autoren den Mehrwert und die Rahmenbedingungen systemischer Ansätze. Zusammenfassend kommen HUMMELBRUNNER et al. zu der Einschätzung, dass eine systemische Herangehensweise an die Evaluation in der Anwendung systemischer Konzepte, Methoden und Haltungen besteht. Diese begründeten keinen neuen ‚systemischen‘ Evaluationsansatz, der sich klar von einer ‚nicht-systemischen‘ Evaluationspraxis abgrenzen ließe. Vielmehr stellten sie eine Erweiterung des Handlungsrepertoires von EvaluatorInnen in komplizierten oder komplexen Situationen dar. Den Autoren zur Folge wirkten systemische Ansätze in schwierigen Situationen, die mit herkömmlichen Methoden nicht – oder nicht ausreichend – analysiert und erklärt werden könnten, etwa bei Widersprüchen, Überraschungen, unerklärlichen oder unerwarteten Phänomenen sinngebend. In derartigen Fällen könnten systemische Ansätze anlassbezogen eingesetzt werden und müssten nicht von vornherein geplant sein.

BUGENHAGEN stellt in ihrem Beitrag *„Objektiv geht nicht mehr.“ – Expert-Innensicht auf systemische Evaluationsansätze in der Entwicklungszusammenarbeit* die Ergebnisse ihrer qualitativen Studie vor. Das wachsende Interesse an systemischen Ansätzen für Evaluationsvorhaben der deutschen Entwicklungszusammenarbeit nahm die Autorin zum Anlass, mit EvaluationsexpertInnen Interviews über die Chancen und Grenzen dieser ‚neuen‘ Herangehensweisen zu führen. Um den Kontext dieser Debatte nachvollziehen zu können, erläutert die Autorin zunächst die in der jüngsten Vergangenheit vollzogenen Veränderungen in der inhaltlichen Ausrichtung und Organisation der deutschen und internationalen Entwicklungszusammenarbeit. Dazu geht BUGENHAGEN auf das sich wandelnde Selbstverständnis der Entwicklungszusammenarbeit, wie die zunehmende Bedeutung einer Makro-Strukturpolitik und die Debatte über Komplexität und Grenzen

der Wirkungsmessung ein. Daran anschließend beantwortet BUGENHAGEN im Rückgriff auf ihre empirische Studie folgende Fragen: Wie werden systemische Ansätze verstanden? Worin besteht Konsens und welche divergierenden Akzentuierungen können identifiziert werden? Welche Rolle kann den EvaluatorInnen zugeschrieben werden und wie lassen sich systemische von nicht-systemischen Ansätzen unterscheiden? Und wie beurteilen die ExpertInnen die Chancen und Grenzen systemischer Evaluationsansätze?

Der Beitrag von SMID mit dem Titel *Lost in Complexity – How to evaluate sustainable development?* stellt ebenfalls Ergebnisse einer empirischen Studie zu systemischen Ansätzen in der Evaluation vor. Der Autor geht der Frage nach, in wie weit systemische Ansätze bei der Evaluation nachhaltiger Entwicklungsprojekte berücksichtigt werden könnten. Im Unterschied zum Beitrag von BUGENHAGEN stellt SMID folglich mit nachhaltiger Entwicklung ein bestimmtes Themenfeld innerhalb der deutschen Entwicklungszusammenarbeit in den Mittelpunkt. Um dies zu ergründen, führt SMID zunächst in die kontroversen Debatten über Nachhaltigkeit ein und arbeitet nachhaltige Entwicklung als komplexes und schwergreifbares Phänomen heraus. Auf der Basis von geführten Experteninterviews mit WissenschaftlerInnen, EvaluatorInnen und AuftraggeberInnen stellt SMID einen breiten Fokus auf Evaluationszwecke, Evaluationsdiskurse, unterschiedliche Positionen und Perspektiven auf systemische Evaluationsansätze vor. Daran anschließend geht SMID auf vier Evaluationsfälle aus der Praxis der deutschen internationalen Zusammenarbeit ein. Er zeigt, wie in den Evaluationen eines Slumentwicklungsprojektes in Nigeria, eines Wasser- und Abwasserprojektes in Äthiopien, eines Projektes zur Ernährungssicherung in den Philippinen und eines Projektes zur ländlichen Entwicklung in Bolivien systemische Ansätze zur Anwendung kamen.

LITERATURVERZEICHNIS

- BAECKER, D. (2008): Zur Evaluation kultureller Projekte. In: Zeitschrift für Evaluation. Heft 1, S. 97–111.
- BATESON, G. (1985): Ökologie des Geistes. Anthropologische, psychologische, biologische und epistemologische Perspektiven. Frankfurt a. M.: Suhrkamp.
- BEYWL, W.; SCHEPP-WINTER, E. (1999): Bundesministerium für Familie, Senioren, Frauen und Jugend (Hg.): Zielfindung und Zielerklärung – Ein Leitfaden. QS Nr. 21. Materialien zur Qualitätssicherung der Kinder- und Jugendhilfe (abrufbar unter URL: <http://www.qs-kompendium.de>, Stand: 26.04.2010).
- BEYWL, W.; SCHEPP-WINTER, E. (2000): Bundesministerium für Familie, Senioren, Frauen und Jugend (Hg.): Zielgeführte Evaluation von Programmen – ein Leitfaden. QS Nr. 29. Materialien zur Qualitätssicherung der Kinder- und Jugendhilfe (abrufbar unter URL: <http://www.qs-kompendium.de>, Stand: 26.04.2010).
- BLANK, S. (2008): Evaluation und Lernprozesse: Eine systemtheoretische Analyse. In: Zeitschrift für Evaluation. Heft 1, S. 75–95.
- BOGUMIL, J.; KUHLMANN, S.; WOLLMANN, H. (2004): Leistungsmessung und -vergleich in Politik und Verwaltung: Konzepte und Praxis. Wiesbaden: VS Verlag für Sozialwissenschaften.
- BUNDESREGIERUNG (2006): Einstieg in das Thema Evaluation. Berlin (abrufbar unter URL: <http://www.verwaltung-innovativ.de>, Stand: 10.06.2011).
- EGNER, H.; RATTER, B. M. W. (2008): Einleitung: Wozu Systemtheorie(n). In: Egner, H.; Ratter, B. M. W.; Dikau, R. (2008) (Hg.): Umwelt als System – System als Umwelt? Systemtheorien auf dem Prüfstand. München: oekom, S. 9–19.
- FETTERMANN, D. M. (1994): Empowerment evaluation. In: Evaluation Practice. 15 (1) 1, S. 1–15.

- GALLA, J.; KOPP, U.; MARTINUZZI, A.; STÖRMER, E. (2008): Programmaktorsaufstellungen – Erste Erfahrungen mit Systemaufstellungen in theoriebasierten Evaluationen. In: Zeitschrift für Evaluation. Jg. 7, Heft 1, S. 35–73.
- GRANZER, D.; WENDT, P.; BERGER, R. (2008): Selbstevaluation in Schulen – Theorie, Praxis und Instrumente. Weinheim: Beltz.
- GUBA, E. G.; LINCOLN, Y. S. (1989): Fourth Generation Evaluation. Thousand Oaks, CA: Sage.
- HOUSE, E. (1993): Professional Evaluation: Social Impact and Political Consequences. Thousands Oaks, CA: Sage.
- HUMMELBRUNNER, R. (2011): Systems thinking and evaluation. In: Evaluation. 17. Heft 4, S. 395–403 (abrufbar unter URL: <http://evi.sagepub.com/content/17/4/395.extract>, Stand: 02.02.2015).
- KEMPFFERT, G.; ROLF, H. G. (2008): Qualität und Evaluation: Ein Leitfaden für pädagogisches Qualitätsmanagement. 4. Überarbeitete und erweiterte Auflage. Weinheim: Beltz.
- KRIZANITS, J. (2005): Die Guten-Tag-Geschichte der systemischen Organisationsberatung. In: Profile. Internationale Zeitschrift für Veränderung, Lernen, Dialog, Heft 10, S. 40–63.
- KROMREY, H. (1995): Evaluation. Empirische Konzepte zur Bewertung von Handlungsprogrammen und Schwierigkeiten ihrer Realisierung. In: Zeitschrift für Sozialisationsforschung und Erziehungssoziologie. Jahrgang 15. Heft 4, S. 313–336.
- KÖNIGSWIESER, R.; HILDEBRAND, M. (2005): Einführung in die systemische Organisationsberatung. 2. Auflage. Heidelberg: Carl-Auer Verlag.
- KÖNIGSWIESER, R.; LANG, E. (2008): Wenn 1 plus 1 mehr als 2 macht. Erfahrung zur Komplementärberatung. In: Zeitschrift für Organisationsentwicklung. Nr. 2/2008, S. 28–39.
- LAY, M.; PAPADOPOULOS, I. (2007): An exploration of fourth generation evaluation in practice. In: Evaluation. Band 13. Heft 4, S. 495–504 (abrufbar unter URL: <http://evi.sagepub.com/cgi/content/refs/13/4/495>, Stand 26.05.2010).
- LUHMANN, N. (1994): Soziale Systeme. Grundriß einer allgemeinen Theorie. Frankfurt a. M.: Suhrkamp.

- MERTENS, D. (2004): Institutionalizing Evaluation in the United States. In: Stockmann, R. (Hg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Sozialwissenschaftliche Evaluationsforschung. Band 1. 2. überarbeitete Auflage. Opladen: Leske und Budrich, S. 45–60.
- PATTON, M. Q. (1987): Evaluation's political inherency: Practical implications for design and use. In: Palumbo, D. J. (Hg.): The politics of program theory. Thousands Oaks. CA: Sage, S. 100–145.
- PATTON, M. Q. (2011): Developmental Evaluation. Applying complexity concepts to enhance innovation and use. New York: The Guilford Press.
- PEEL, D. (2003): Town Centre Management: Multi-stakeholder Evaluation. Increasing the Sensitivity of Paradigm Choice. In: Planning Theory & Practice. Band 4. Heft 2, S. 147–164.
- ROLFES, M.; WILHELM, J. L. (2014): Evaluationspraxis und Evaluationsforschung im Kontext der Stadt- und Regionalentwicklung. In: Böttcher, W. et al. (Hg.): Evaluation in Deutschland und Österreich. Stand und Entwicklungsperspektiven in den Arbeitsfeldern der DeGEval – Gesellschaft für Evaluation. Waxmann. Münster, S. 21–35.
- ROMMELSBACHER, T. (2001): Ansprüche an Evaluation aus Sicht von Politik und Wissenschaft. In: Institut für Landes- und Stadtentwicklungsforschung des Landes Nordrhein-Westfalen (ILS) (Hg.): Monitoring und Controlling in Stadtteilen mit besonderem Erneuerungsbedarf. Reader zum Workshop in Hamm. Dortmund, S. 6–8.
- SEDLACEK, P. (2004): Evaluation in der Stadt- und Regionalentwicklung. Herausforderung für Wissenschaft und Praxis. In: Sedlacek, P. (Hg.): Evaluation in der Stadt- und Regionalentwicklung. Stadtforschung aktuell. Band 90. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 11–26.
- SPENCER BROWN, G. (1977): Laws of Form. New York.
- STAKE, R. E. (1983): The case study method in social inquiry. In: Madaus, G. F.; Scriven, M.; Stufflebeam, D. L. (Hg.): Evaluation Models. Boston: Kluwer-Nijhoff, S. 279–286.
- STOCKMANN, R. (2004) (Hg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Sozial-wissenschaftliche Evaluationsforschung Band 1. 2. überarbeitete Auflage. Opladen: Leske und Budrich.

- STOCKMANN, R. (2006): Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Sozialwissenschaftliche Evaluationsforschung. Band 5. Münster: Waxmann.
- STOCKMANN, R.; MEYER, W. (2010): Evaluation. Eine Einführung. Opladen: Barbara Budrich.
- ULRICH, S.; WENZEL, F. M. (2003): Partizipative Evaluation. Ein Konzept für die politische Bildung. Gütersloh: Verlag Bertelsmann Stiftung.
- V. AMELN, F. (2004): Konstruktivismus. Die Grundlagen systemischer Therapie, Beratung und Bildungsarbeit. Tübingen: UTB.
- V. SCHLIPPE, A.; SCHWEITZER, J. (2007): Lehrbuch der systemischen Therapie und Beratung. 10. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- VARGA VON KIBED, M.; SPARRER, I. (2011): Ganz im Gegenteil. Tetralemmaarbeit und andere Grundformen systemischer Strukturaufstellungen – für Querdenker und solche, die es werden wollen. 7. Auflage. Heidelberg: Carl-Auer Verlag.
- VOGD, W. (2005): Systemtheorie und rekonstruktive Sozialforschung. Eine empirische Versöhnung unterschiedlicher theoretischer Perspektiven. Opladen: Barbara Budrich.
- WATZLAWICK, P.; BEAVIN, J. H.; JACKSON, D. D. (2011): Menschliche Kommunikation. Formen Störungen Paradoxien. 12. Auflage. Bern: Verlag Hans Huber.
- WILHELM, JAN L. (2012): Wozu Evaluation? Organisationssysteme bewerten Stadtteilmförderung mit Kalkül. Potsdam: Universitätsverlag.
- WILLIAMS, B.; IMAM, I. (Hg.) (2006): Systems concepts in evaluation. An expert anthology. American Evaluation Association/Edge Press: Point Reyes (abrufbar unter URL: <http://preval.org/files/Kellogg%20enfoque%20sistematico%20en%20evaluacion.pdf>, Stand: 02.02.2015).
- WOLLMANN, H. (2005): Evaluierung von Verwaltungsmodernisierung. In: Blanke, B.; v. Bandemer, S.; Nullmeier, F.; Wewer, G. (Hg.): Handbuch zur Verwaltungsreform. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 502–510.

SYSTEMISCHE ANSÄTZE IN DER EVALUATION¹

Richard Hummelbrunner | Bernward Causemann | Theo Mutter | Michaela Raab

¹ Mit dieser Publikation stellt die AG „Systemische Ansätze in der Evaluation“ des AK Entwicklungspolitik in der DeGEval ihr Arbeitsergebnis vor. In 19 Sitzungen von 2008 bis 2012 wurde zunächst eine thematisch-fachliche Grundlage erarbeitet; im Anschluss daran lud die AG im Mai 2011 auf der Frühjahrstagung des AK Entwicklungspolitik die Fachöffentlichkeit zur vertiefenden Diskussion systemischer Ansätze in der Evaluation ein. In der abschließenden Phase wurde in der AG das hier vorgelegte Diskussionspapier erarbeitet, das einen kurzen Überblick zum Thema gibt und zugleich als Einstieg in die Thematik dienen soll. Dieses Papier wurde auf der Frühjahrstagung des AK Entwicklungspolitik im Juni 2013 diskutiert und als Arbeitspapier des AK Entwicklungspolitik in der DeGEval zur Publikation freigegeben.

1 EINLEITUNG

Systemisches Denken und Evaluation sind breit gefächerte Bereiche, jeder mit einer Reihe von Theorien, Schulen und Herangehensweisen. Sie haben sich bisher weitgehend unabhängig voneinander entwickelt. Erst seit einigen Jahren gibt es Bestrebungen, sie gezielter miteinander zu verbinden. Dabei haben die bisherigen Überlegungen gezeigt, dass es weder möglich noch sinnvoll ist, einen spezifischen Evaluationstypus als Verbindung dieser beiden Bereiche einzuführen. Es gibt daher keine ‚Systemische Evaluation‘ mit klar abgrenzbaren Unterscheidungsmerkmalen zu einer ‚nicht-systemischen‘ Evaluationspraxis. Die Anwendung systemischer Ansätze steht also in keinerlei Widerspruch zu den professionellen Standards der DeGEval und den DAC/OECD Kriterien.

Systemisches Denken ist eine spezifische Art die Realität zu begreifen und zu erklären. Während früher Systeme als real existierende Objekte angesehen wurden, werden sie heute zumeist als mentale Konstrukte zur Erklärung der Wirklichkeit aufgefasst. Um eine Situation als System zu erfassen, muss man sie einem Zweck zuordnen und von der jeweiligen Umwelt abgrenzen. Ferner sollten die maßgeblichen Elemente und ihre Beziehungen identifiziert werden. Diese Form der Betrachtung ist beobachterabhängig und ist besonders hilfreich für die Arbeit mit sozialen Systemen (z. B. Individuen, Gruppen, Organisationen, Netzwerken).

Innerhalb des Systemischen Denkens haben sich eine Reihe von Schulen und Ausrichtungen herausgebildet, die sich in Bezug auf Fokus und Methoden unterscheiden. Diese Differenzierung ist v. a. im anglo-amerikanischen Raum gebräuchlich (z. B. *Cybernetics*, *System Dynamics*, *Soft Systems*, *Complex Adaptive Systems*, *Critical Systems*). Was im deutschen Sprachraum unter ‚systemisch‘ verstanden wird, ist eine im Laufe der Zeit entstandene Ansammlung von Elementen dieser verschiedenen Schulen. Da AnwenderInnen in der Entwicklungszusammenarbeit unterschiedliche systemische Traditionen vertreten, wurde diesem Papier bewusst nicht (nur) das im deutschen Sprachraum geläufige systemische Verständnis zu Grunde gelegt. Der vorliegende Text beruht vielmehr auf jenem integrativen Verständnis, das sich unter internationalen Systemdenkern als Verbindung der verschiedenen Ausrichtungen herausgebildet hat. Mit diesem breiteren Ansatz soll sowohl der Einstieg für Personen mit geringen systemischen Vorkenntnissen als auch die Anschlussfähigkeit im internationalen Umfeld erleichtert werden.

Da Systemisches Denken nicht auf einem homogenen, abgeschlossenen Theoriegebäude beruht, gibt es auch keine einheitliche Definition von *System* und *systemisch*. Und manches von dem, was als systemisch angesehen wird, ist in der

Entwicklungszusammenarbeit bereits bekannt (und wird oft nicht so bezeichnet²), oder findet sich in anderen Evaluationsansätzen wieder (z. B. in der partizipatorischen oder anwendungsorientierten Evaluation). Diese mangelnde Trennschärfe sollte jedoch weder dazu führen, den Begriff systemisch in Bezug auf Evaluation fallenzulassen, noch dazu verleiten, alles mit diesem Etikett zu versehen. Vielmehr geht es darum, die wesentlichen Merkmale zu definieren und daraus – im Sinne einer Qualitätssicherung – Kriterien für die Praxis abzuleiten und Unterschiede zu anderen Evaluationsansätzen herauszuarbeiten, wo dies möglich und sinnvoll ist.

Das zentrale Merkmal einer systemischen Herangehensweise ist die Anwendung systemischer Konzepte, Methoden und Haltungen. Die Basis bilden drei generische systemische *Konzepte* (Beziehungen, Perspektiven, Systemgrenze), die den gemeinsamen Nenner der verschiedenen Ausrichtungen des systemischen Denkens bilden. Diesen können jeweils eine Reihe von *Methoden* zugeordnet werden, für deren Anwendung eine spezifische systemische *Haltung* der an den Evaluationen beteiligten Personen erforderlich ist. Systemische Ansätze sind insbesondere dann geeignet, wenn Evaluationen mit Komplexität konfrontiert sind, d. h. mit Situationen, die durch viele Elemente mit verschiedenartigen Beziehungen gekennzeichnet sind und die von den Beteiligten unterschiedlich wahrgenommen werden.

Forschungsarbeiten der letzten Jahre erlauben eine Differenzierung und Operationalisierung des Begriffes Komplexität in Bezug auf Evaluation. Tabelle 1 fasst die drei Zustände (Domänen), ihre wesentlichen Merkmale und die dafür geeigneten Monitoring- und Evaluationsansätze (M&E) zusammen.

Bei den Domänen handelt es sich um idealtypische Zustände, die zwar nach bestimmten Merkmalen unterschieden werden können, zumeist aber in kombinierter Form auftreten. Wichtig ist ferner, dass nicht eine gesamte Situation (z. B. ein Vorhaben) einer Kategorie zugeordnet werden muss, sondern nur bestimmte Teile oder Aspekte (z. B. Komponenten eines Vorhabens).

Zwar gibt es auch andere Systematisierungen, doch dominiert die oben vorgestellte derzeit in der internationalen Fachdiskussion. Sie erlaubt eine grobe Einschätzung der Sinnhaftigkeit einer Anwendung systemischer Ansätze bei M&E: In Situationen, die der Domäne des Einfachen zugeordnet werden können, wird Systemisches Denken kaum erforderlich sein. Hingegen sind systemische Ansätze bei der Arbeit mit komplizierten Situationen hilfreich: zum Beispiel für die Visualisierung vernetzter oder nicht-linearer Beziehungen, die Analyse und Bearbeitung unterschiedlicher Perspektiven, die Abgrenzung von System (z. B. Vorhaben) und Umwelt oder die Reflexion über deren Wechselbeziehungen bzw. über die Konsequenzen derartiger Abgrenzungen. In komplexen Situationen sind systemische Ansätze unabdingbar, denn nur sie bieten Methoden zur adäquaten Erfassung von Mustern und Beziehungen zwischen den Elementen bzw. zwischen den AkteurInnen an.

² Es treten häufig Vermischungen mit partizipativen Ansätzen oder der Aktionsforschung auf, die entweder auf systemischem Denken beruhen oder mit ähnlichen Konzepten und Herangehensweisen operieren.

Domäne	Charakteristika	Ursache-Wirkung	M&E Ansätze
Einfach ,Bekannt'	<ul style="list-style-type: none"> - Hohe Sicherheit und Zustimmung der Beteiligten - Bekannte richtige Antwort - ‚Best practice‘ Rezepte 	<ul style="list-style-type: none"> - Klar, vorhersehbar und kontrollierbar 	<ul style="list-style-type: none"> - Überprüfung von ‚best practice‘ - ‚Summative‘ Evaluation mit linearen Modellen
Kompliziert ,Erfassbar'	<ul style="list-style-type: none"> - Einige Unsicherheiten und unterschiedliche Ansichten - ‚Good practice‘ identifizierbar - Erfordert Analysen, Koordination und Expertise 	<ul style="list-style-type: none"> - Plan- und vorhersagbar (aber erfordert hohes Fachwissen) - Kontextabhängig - Alternative Pfade 	<ul style="list-style-type: none"> - Überprüfung tatsächlicher Praxis (Monitoring & formative Evaluation) - ‚Summative‘ Evaluation auf Basis einer überprüfbarer Theorie
Komplex ,Nicht erfassbar'	<ul style="list-style-type: none"> - Hohe Unsicherheit, Widersprüche - Jede Situation ist einzigartig - Erfordert die Beobachtung von Beziehungen und Mustern 	<ul style="list-style-type: none"> - Klarheit nur im Rückblick - Abhängigkeit von der Ausgangslage 	<ul style="list-style-type: none"> - Dokumentation der Ausgangslage - Erfassung von Mustern - Begleitendes, rasches Feedback

Tab. 1: Domänen von Situationen

Quellen: GLOUBERMAN und ZIMMERMAN (2002), KURTZ und SNOWDEN (2003), PATTON (2010), ROGERS (2008)

Im folgenden Kapitel werden die Hauptbestandteile systemischer Ansätze (Konzepte, Methoden, Haltungen) beschrieben. Das zweite Kapitel betrachtet den Evaluationsprozess aus systemischer Sicht und zeigt die wesentlichen Implikationen auf (z. B. für die Rolle von EvaluatorInnen). Das dritte Kapitel stellt systemische Herangehensweisen zur Analyse von Wirkungen und Kausalität als zentrales Element von Evaluationen vor. Das vierte Kapitel geht auf die erforderlichen Rahmenbedingungen für die Anwendung systemischer Ansätze und deren Mehrwert ein.

2 DIE HAUPTBESTANDTEILE SYSTEMISCHER ANSÄTZE IN EVALUATIONEN

Die Essenz systemischen Denkens kann in drei Konzepten zusammengefasst werden, die allen systemischen Schulen gemeinsam sind und deren historische Entwicklung widerspiegeln. Diese Konzepte sind zwar im Bereich der Evaluation durchaus geläufig, bieten jedoch im Rahmen von systemischem Denken einen besonderen Fokus und eine Reihe von spezifischen Methoden an. Nachfolgend werden diese drei generischen Systemkonzepte kurz beschrieben und exemplarisch anhand einiger für Evaluationen besonders geeigneter Methoden vorgestellt (siehe dazu insbesondere WILLIAMS, HUMMELBRUNNER 2011).

2.1 Beziehungen

Beziehungen stellen das älteste und wahrscheinlich bekannteste Systemkonzept dar. Es befasst sich mit der Verknüpfung der Elemente eines Systems und den daraus erwachsenden Konsequenzen. Dabei wird einerseits die Qualität von Beziehungen betrachtet, insbesondere deren Sensitivität gegenüber Änderungen bei anderen Elementen und Beziehungen im System oder in der Systemumwelt. Das Ausmaß dieser Sensitivität wird mit dem Begriff der Kopplung erfasst: Je mehr zwei Systemelemente einander beeinflussen, desto stärker sind sie gekoppelt. Andererseits geht es um die Dynamik und den Verlauf von Beziehungen; v. a. um nicht-lineare Aspekte, die durch Rück- und Wechselwirkungen (Feedback) zwischen Elementen entstehen und im Zeitablauf zu neuen Mustern, Strukturen und Eigenschaften führen können. Letzteres wird als ‚Emergenz‘ bezeichnet.

Systemische Methoden erlauben die dynamische Modellierung von Beziehungen, wobei je nach Zweck zwischen zwei grundlegenden Typen unterschieden werden kann: Methoden zur Gewinnung von Einsichten in die Funktionsweise von Systemen oder zur Simulation bzw. Prognose künftiger Zustände. Bei den im Anschluss beschriebenen Methoden ist jeweils angeführt, für welche Zwecke sie eingesetzt werden können.

KYBERNETISCHE METHODEN

Die Kybernetik, die sich mit der Regelung vernetzter Abläufe befasst, hat eine auf Kreisen basierende Visualisierungsform entwickelt. Sie baut auf dem Konzept des Regelkreises auf, das aufzeigt, wie sich Handlungen wechselseitig verstärken oder ausgleichen. Kybernetische Methoden geben in die Funktionsweise von Systemen Einblick und sind besonders für die Beschreibung von Wechselbeziehungen und Veränderungsmustern geeignet. Durch den Fokus auf die grundlegenden Strukturen (statt auf einzelne Ereignisse) können auch geeignete Ansatzpunkte für deren Veränderung identifiziert werden. Die Modelle sind in der Regel wenig formalisiert und arbeiten mit qualitativen Größen oder verbalen Beschreibungen.

Die Visualisierung von Beziehungen ist mit einfachen Hilfsmitteln möglich (Flipchart, Kärtchen), kann aber auch mittels spezieller Software erfolgen. Die Elemente werden durch Pfeile verbunden, deren Spitze die Bewegungsrichtung der Beziehung anzeigt (Zunahme, Abnahme). Damit können z. B. die Einflussfaktoren auf eine Situation oder die Ursachen für ein Problem dargestellt werden. Bei Wirkungsdiagrammen werden die Beziehungen zu geschlossenen Regelkreisen verbunden, die entweder verstärkend oder ausgleichend wirken.

Kybernetische Methoden fördern das Verständnis für Zusammenhänge und erlauben Aussagen über Zustände und Beziehungsmuster innerhalb eines Systems und ggf. zwischen System und Umwelt. Sie eignen sich besonders für die Situations- oder Problemanalyse, da sie die Erfassung dynamischer Komplexität und der Tiefenstruktur von beobachtbaren Phänomenen (Symptomen) ermöglichen.

BEISPIEL: WIRKUNGSDIAGRAMM ZUR WELTBEVÖLKERUNGS- ENTWICKLUNG

Die Erstellung eines Wirkungsdiagramms erfolgt in drei Schritten:

1. Ermittlung der Elemente

Zunächst werden jene Faktoren ermittelt, die für eine Situation von Relevanz sind. Diese Faktoren müssen nicht zahlenmäßig festlegbar sein, es muss aber möglich sein, ihre Bewegungsrichtung anzugeben (Zunahme, Abnahme). Um die Komplexität auf ein bearbeitbares Maß zu reduzieren, muss eine Auswahl getroffen werden, die sich nach der jeweiligen Aufgabenstellung richtet. Für die Bevölkerungsentwicklung werden in diesem Fall vier Faktoren ausgewählt: Geburten und Todesfälle sowie die jeweiligen Veränderungsdaten. Da es sich um eine vereinfachte Darstellung handelt, werden andere Einflussfaktoren, wie z. B. die Zu-/Abwanderung, nicht berücksichtigt.

2. Erfassung der Beziehungen

Die Zusammenhänge zwischen den Elementen werden dargestellt, indem die Elemente durch Pfeile verbunden werden, die die Richtung der Beeinflussung anzeigen. Ein Vorzeichen am Pfeilende gibt die Art des Einflusses an. Die Qualität der Beziehungen kann bei Bedarf noch weiter analysiert werden, z. B. deren zeitlicher Verlauf, die Kontinuität des Verlaufs oder die zeitlichen Verzögerungen.

3. Verknüpfung zu Regelkreisen

Anschließend werden die Elemente zu zusammengehörenden Kreisläufen verbunden, wodurch ein aus Regelkreisen zusammengesetztes Wirkungsnetz entsteht. Dabei ist wichtig, dass alle Elemente erfasst werden und geschlossene Wirkungsketten entstehen. Elemente ohne offenkundige Verbindungen können entweder weggelassen werden oder es muss nach den noch fehlenden Bindegliedern gesucht werden.

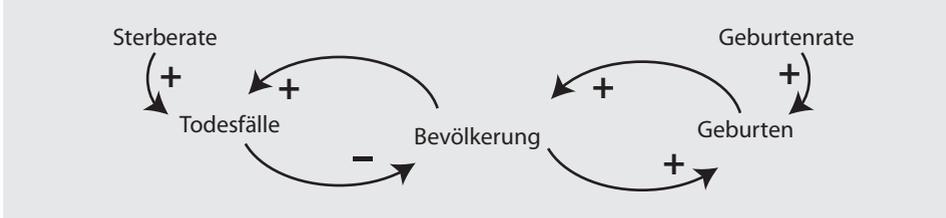


Abb. 1: Wirkungsdiagramm; Quelle: eigene Darstellung

SYSTEM-DYNAMISCHE METHODEN (SYSTEM DYNAMICS)

Sie beruhen ebenfalls auf dem Konzept des Regelkreises, arbeiten aber ausschließlich mit quantitativen Größen und erlauben damit eine mathematische Berechnung und Prognose von Systemzuständen. System-dynamische Methoden entsprechen strukturell den Wirkungsdiagrammen (und können auch von diesen abgeleitet werden). Sie unterscheiden sich allerdings in der bildlichen Darstellung und sind spezifischer, da sie genau festlegen, wie die einzelnen Verbindungen funktionieren. Die Simulation erlaubt Aussagen über die von einer Systemstruktur erzeugte Dynamik und die Identifizierung wirksamer Hebel für Veränderungen. Dies ist vor allem in Situationen hilfreich, in denen die menschliche Intuition nicht ausreicht, z. B. weil die Konsequenzen von Handlungen weder sofort spürbar noch eindeutig sind.

Ausgangspunkt von System Dynamics ist die Überlegung, dass es in jedem System Elemente gibt, die zu einem bestimmten Zeitpunkt einen Wert haben (Bestand), der sich im Lauf der Zeit durch Zu- oder Abflüsse verändern kann. Die Beziehung zwischen diesen Größen wird in Flussdiagrammen dargestellt, wobei die Pfeilrichtung angibt, welche Flussgröße auf welchen Zustand einwirkt. Die

Beziehungen zwischen den Größen müssen quantifizierbar sein und werden durch stark formalisierte Beziehungen abgebildet (z. B. mathematische Funktionen, Gleichungen). Auf dieser Basis können mittels spezieller Software Simulationsmodelle erstellt werden, die das dynamische Verhalten von Systemen berechnen und die Prognose künftiger Zustände erlauben (vgl. O'CONNOR, MCDERMOTT 2000; WILLIAMS, IMAM 2007).

**BEISPIEL:
FLUSSDIAGRAMM ZUR JÄHRLICHEN BEVÖLKERUNGS-
ENTWICKLUNG**

Bei der Erstellung von Flussdiagrammen wird am besten von einer (oder mehreren) Grundbeziehung(en) ausgegangen. Diese wird (werden) sukzessive durch das Hinzufügen weiterer Faktoren erweitert, die auf die Flussgrößen einwirken und in Form von Kausalschleifen dargestellt werden. Flussdiagramme bestehen daher – analog zu Wirkungsdiagrammen – aus miteinander vernetzten Regelkreisen. Auch in diesem Modell wird zwischen verstärkenden und ausgleichenden Rückkopplungen unterschieden.

Das oben beschriebene Wirkungsdiagramm zur Bevölkerungsentwicklung kann auch als Flussdiagramm dargestellt werden. Die Bevölkerungszahl sowie die Geburten- und Sterberaten werden dabei graphisch als Bestandsgröße (□) und die jährlichen Geburten und Todesfälle als Flussgröße (⊖) dargestellt. Sie werden anschließend durch Pfeile verbunden, wobei die Pfeilrichtung angibt, welche Flussgröße auf welchen Zustand einwirkt. Eine ‚Wolke‘ (☁) symbolisiert die Grenze des betrachteten Systems.

Um die funktionalen Beziehungen detaillierter darzustellen, können Hilfsvariablen verwendet werden. Besondere Bedeutung kommt den Verzögerungsvariablen zu. Sie dienen dazu, Zeitverzögerungen in das System einzuführen oder Informationen über frühere Zustände einzelner Variablen zu erhalten.

Die einzelnen Variablen und deren Beziehungen werden in der Folge quantifiziert und erlauben damit eine Berechnung der Veränderung. Für jede kausale Beziehung zwischen Variablen muss daher eine mathematische Funktion gefunden werden, die entweder lineare oder nicht-lineare Beziehungen abbildet (letztere können oft nur geschätzt werden).

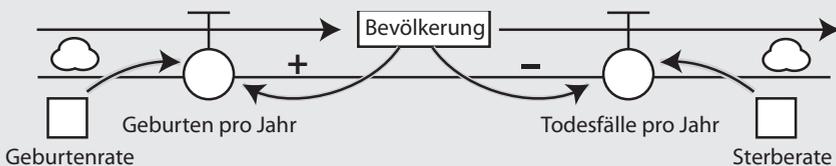


Abb. 2: Flussdiagramm; Quelle: eigene Darstellung

SOZIALE NETZWERKANALYSE (SNA)

Die SNA basiert auf soziologischen Theorien, die soziale Phänomene durch die Interaktion von Individuen erklären und das (Kommunikations-)Verhalten von (Klein-)Gruppen untersuchen. Durch deren Verbindung mit der Graphentheorie und die Verwendung computergestützter mathematischer Methoden können auch größere soziale Systeme analysiert und eindrucksvoll visualisiert werden. Im Gegensatz zu einfacheren Darstellungen von Beziehungsmustern (z. B. Akteurslandkarte, Venn-Diagramme) erlaubt die Soziale Netzwerkanalyse eine vertiefende Analyse und die Simulation von Veränderungen.

Die Beziehungen in sozialen Systemen werden als Netzwerk konzipiert und durch drei miteinander kombinierbare Techniken analysiert: (1) Daten und Informationen über die Beziehungen der einzelnen Akteure werden erhoben und zunächst in Matrixform zusammengefasst, d. h. als Markierungen oder Werte an den Kreuzungspunkten der Elemente (Spalten und Zeilen). (2) Mittels spezieller Software werden (daraus Netzwerkdiagramme erstellt, um Beziehungen zu visualisieren und nach unterschiedlichen Gesichtspunkten zu strukturieren. (3) Beziehungen können auch quantitativ analysiert und ihre strukturellen Charakteristika mittels Kenngrößen berechnet werden (z. B. Dichte eines Netzwerkes, Zentralität von Akteuren). Mit dieser Methode können sowohl das soziale System als Ganzes als auch die Positionen einzelner Akteure untersucht werden. Zudem erlaubt sie die Simulation von Veränderungen im Netzwerk (vgl. JANSEN 2006). Die Abbildung des Kommunikationsgefüges nutzt das oft nur implizite Wissen über Beziehungen und kann Reflexionen zu deren besserer Gestaltung anregen.

BEISPIEL:

NETZWERKDIAGRAMM DER BEZIEHUNG VON NRO IN GHANA

Zunächst wurden die Beziehungen zwischen den NRO erhoben und in Form einer (Beziehungs-)Matrix dargestellt. Dabei wurden in den Zeilen (horizontal) und Spalten (vertikal) jeweils die gleichen Akteure aufgelistet. Jede Zelle (Kreuzungspunkt) enthält die Zahl der Beziehungen zwischen den jeweiligen NRO. Die Art der Beziehung und der Beobachtungszeitraum wurden vor der Erhebung spezifiziert Die Informationen wurden anschließend mit einer SNA Software in ein Netzwerkdiagramm umgewandelt. In einem derartigen Bild wird die Struktur der Beziehungen besser sichtbar als in einer Matrixdarstellung. Komplexe Information wird verdichtet; signifikante Muster können leichter erkannt werden.

Das Diagramm (Abb. 3) gibt Aufschluss darüber, welche Akteure im System zentral (z. B. WILDAF) oder peripher (z. B. FIDA) angesiedelt sind; die Größe der Kreise spiegelt die Anzahl der Beziehungen eines Akteurs mit allen anderen wider.

Die Pfeile symbolisieren die Intensität der Beziehungen sowie deren Richtung (einseitig oder wechselseitig). Das Bild gibt weiterhin Aufschluss über die Dichte des Netzwerks und die Existenz von eng verflochtenen Untergruppen und von relativ isolierten Akteuren (z. B. oben links). Dieses Diagramm wurde zu einer gemeinsamen Auswertung und Reflexion des Netzwerks genutzt: Die beteiligten Akteure haben es vor dem Hintergrund ihrer eigenen Erfahrungen interpretiert und Überlegungen zur Verbesserung des Netzwerks in Bezug auf Effizienz und Stabilität angestellt.³

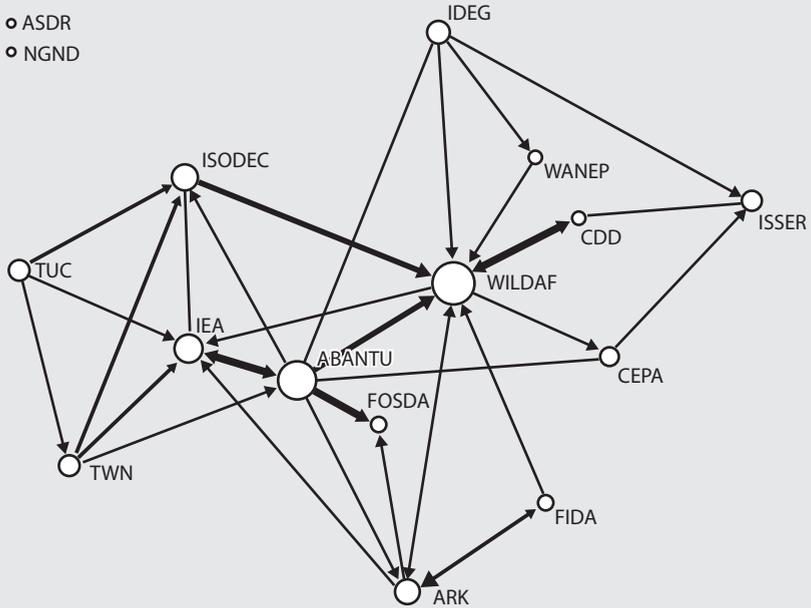


Abb. 3: Netzwerkdiagramm; Quelle: <http://mande.co.uk/special-issues/network-models/>

³ Quelle: <http://mande.co.uk/special-issues/network-models/>. Eine Liste von Softwares für die ‚Soziale Netzwerkanalyse‘ findet sich auf <http://mande.co.uk/2008/lists/social-network-analysis-software-a-list/> (07.09.2012).

2.2 Perspektiven

Systemische Methoden dienen einerseits dazu, Situationen unter Berücksichtigung unterschiedlicher Perspektiven zu modellieren. Andererseits ermöglichen sie den Austausch zwischen den Akteuren, ein besseres Verständnis der jeweiligen Sichtweisen sowie die Überwindung von Unterschieden oder die Gewinnung neuer Einsichten bzw. Handlungsoptionen.

Die Darstellung und Bewertung von Beziehungen stellen keine objektiven Festlegungen dar; ihre Ermittlung ist subjektiv und erfolgt daher immer aus einer bestimmten (Beobachtungs-)Perspektive. Dabei kann es sich sowohl um die Sichtweisen von Individuen, als auch um die sozialer Systeme (z. B. Organisationen) handeln. Systemische Ansätze verstehen unter ‚Perspektive‘ nicht nur die Sichtweisen von Beteiligten (*Stakeholdern*), die wiederum auf Interessen, Werten oder Gewohnheiten beruhen. Es können auch die Anliegen (*stakes*) herausgearbeitet werden, nach denen eine Situation betrachtet werden kann. Letztere geben Aufschluss über sachlich begründete Motive und Verhaltensweisen.

Diese Unterscheidung erlaubt eine für das Verständnis und die Veränderung sozialer Systeme entscheidende Differenzierung: Stakeholder haben in der Regel mehr als ein Anliegen; unterschiedliche Stakeholder können die gleichen Anliegen teilen. In dem unten angeführten Beispiel zu *Soft Systems Methodology* (SSM) stehen z. B. Anliegen (Evaluationsansprüche) im Vordergrund der Betrachtung. Zusätzlich kann aber auch herausgearbeitet werden, welche der Beteiligten diese Anliegen (in welchem Maß) vertreten. Dadurch werden die mit den Anliegen verknüpften sozialen Systeme identifiziert. Dies kann z. B. dazu genutzt werden, das Maß der Zustimmung durch die Beteiligten zu einem ermittelten Kompromissvorschlag auszuloten oder die Bildung von Allianzen abzuschätzen.

SOFT SYSTEMS METHODOLOGY (SSM)

Dieser Ansatz ist Ende der 1960er Jahre in Großbritannien als Reaktion auf die Erkenntnis entstanden, dass ‚Hard Systems‘ (wie z. B. *Systems Engineering*, *Operational Research*) für die Lösung von (Management-)Problemen vielfach unzureichend waren. Der wesentliche Grund dafür lag in der Vernachlässigung der Perspektiven, aus denen heraus eine Situation betrachtet werden konnte. SSM gilt daher als Prototyp für die Modellierung einer Situation aus unterschiedlichen Perspektiven.

Bei SSM wird eine Situation aus mehreren Perspektiven betrachtet, um unterschiedliche Sichtweisen zu ermitteln und Ansatzpunkte für Veränderungen zu identifizieren. Zunächst wird herausgearbeitet, welche Perspektiven für die Betrachtung

tung einer Situation relevant sind. Die Situation wird dann aus den unterschiedlichen Perspektiven heraus kurz als System beschrieben. Anschließend werden die Perspektiven (als verbale und graphische Modelle) sowohl untereinander als auch mit der Situation (aus der Perspektive des ‚externen‘ Beobachters) verglichen. Dadurch wird ein besseres Verständnis der Perspektiven, ihrer logischen Konsequenzen und ihres Erklärungswerts gewonnen (vgl. REYNOLDS, HOWELL 2010).

BEISPIEL: KLÄRUNG VON EVALUATIONSANSPRÜCHEN MITTELS SSM

Die Evaluation eines als problematisch empfundenen Projekts der Entwicklungszusammenarbeit (EZ) war mit unterschiedlichen Ansprüchen konfrontiert. In einem Klärungsworkshop mit den Beteiligten wurden diese Ansprüche zu mehreren Perspektiven verdichtet und kurz charakterisiert. Dazu zählten:

- » **Gelbe Karte:** Die Evaluation dient dazu, dem Geldgeber eine (Früh-) Warnung über das Funktionieren des untersuchten Projekts zu geben.
- » **Besseres Management:** Die Evaluation dient dazu, Verbesserungen bei den durchzuführenden Prozessen zu erarbeiten.
- » **Vertrauensbildung:** Die Evaluation dient dazu, Probleme auszuräumen und das Vertrauen in die Leistungsfähigkeit des Projekts wiederherzustellen.
- » **Test:** Die Evaluation dient dazu, die Eignung bestimmter Akteure für die Arbeit mit der Community (im Community Development-Projekt) abzuklären.
- » **Policy-Beitrag:** Die Evaluation dient dazu, Erfahrungen mit dem gewählten Handlungsansatz für andere zu gewinnen.

Für jede dieser Perspektiven wurden die Grundanforderungen herausgearbeitet und zu einem Evaluationsmodell verdichtet (v. a. Leitfragen, Methodik). Anschließend wurden diese Modelle sowohl untereinander (v. a. im Blick auf gemeinsame oder sich ausschließende Elemente) als auch mit der ‚realen‘ Situation (Auftragsrahmen, Zeit- und Kostenrahmen) verglichen. Daraus wurde ein Vorschlag für die Ausrichtung der Evaluation abgeleitet, der mehrere Perspektiven bzw. Elemente berücksichtigte. Dadurch konnte die Akzeptanz der Evaluation sowohl für die Geldgeber als auch die Evaluierten deutlich verbessert werden.

ZIRKULÄRE DIALOGE

Diese Dialogform stellt ein Lernsystem auf Basis verschiedener Perspektiven dar. Während Diskussionen oder Debatten der Auseinandersetzung und der Markierung von Unterschieden dienen, nutzen Dialoge die Intelligenz einer Gruppe, um zu gemeinsamen Einsichten zu gelangen, die sich den einzelnen Teilnehmenden schwer erschließen würden. Lernschritte werden durch gezielte Perspektivenwechsel ausgelöst, die die Wahrnehmung erweitern und intensivieren. Zirkuläre Dialoge kommen aus den Traditionen der Rollenspiele und der systemischen Interventionstechniken.

Bei einem zirkulären Dialog wird in einem Gruppenprozess eine Situation aus (zumindest) drei Perspektiven betrachtet. Diese Perspektiven können entweder *Stakeholdern* zugeordnet werden oder – wie im unten angeführten Beispiel – funktional zugeschriebene Rollen repräsentieren. Durch die strukturierte Abfolge von Befragung und Beobachtung sollen eine direkte Diskussion unter den Teilnehmenden unterbunden und deren Aufmerksamkeit erhöht werden. Unterschiedliche Sichtweisen werden dadurch besser verstanden – oder ein neues, gemeinsames Verständnis erreicht. Damit können auch Blockaden infolge verhärteter (Rollen-) Zuschreibungen gelöst sowie die Kreativität und Offenheit für neue Lösungen gefördert werden (vgl. BAUMFELD, HUMMELBRUNNER, LUKESCH 2009).

BEISPIEL:

ABLAUF EINES ZIRKULÄREN DIALOGS ZUR REFLEXION VON EVALUATIONSERGEBNISSEN

Zu Beginn präsentieren die EvaluatorInnen einige ausgewählte Ergebnisse oder vorläufige Bewertungen. Im Anschluss werden die Teilnehmenden in drei Gruppen unterteilt, die drei für die Reflexion des Präsentierten sinnvolle Rollen repräsentieren: Euphoriker, Kritiker und Pragmatiker. Nach einer Einstiegsphase, in der sich die Gruppen auf ihre Rolle vorbereiten, verläuft der Dialog in (zumindest) sechs Schritten und wird von einem/r ModeratorIn geleitet:

Zunächst erläutern die Euphoriker jene Argumente, die aus ihrer Sicht die präsentierten Ergebnisse positiv untermauern. Anschließend erläutern die Kritiker, welche Aspekte sie skeptisch sehen oder hinterfragen, und begründen warum. Die Pragmatiker beobachten den Austausch zunächst, wägen die Argumente ab und zeigen Lösungsmöglichkeiten auf (eventuell angeleitet durch spezifische Fragen oder Anliegen der EvaluatorInnen).

Anschließend haben Euphoriker und Kritiker jeweils die Möglichkeit, auf die Vorschläge der Pragmatiker zu reagieren. Die Pragmatiker reflek-

tieren über diese zweite Dialogrunde und modifizieren gegebenenfalls ihre Vorschläge oder bringen neue ein.

Abschließend kommentieren die EvaluatorInnen den gesamten Dialog, insbesondere welche Erkenntnisse oder Anregungen sie daraus für die weitere Bearbeitung gewonnen haben. Je nach Größe der Gesamtgruppe benötigt dieser Dialog eine Gesamtdauer von 45 bis 90 Minuten. Bei Bedarf kann auch eine dritte Dialogrunde durchgeführt werden.⁴

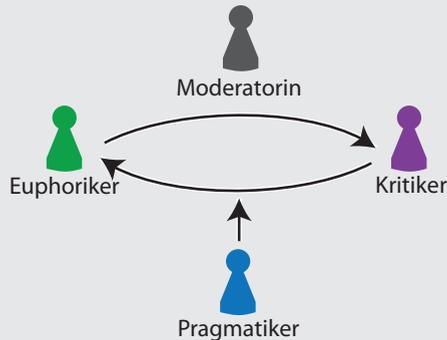


Abb. 4: Zirkulärer Dialog; Quelle: eigene Darstellung

DIALEKTISCHE UNTERSUCHUNGSMETHODEN

Dialektik ist eine Methode der Gesprächsführung, mit der Erkenntnisse durch die Auslotung gegensätzlicher Standpunkte gewonnen werden. Dialektik hat sowohl im ‚westlichen‘ als auch im ‚östlichen‘ Denken eine lange Tradition. Dabei werden gezielt Unterschiede (statt Gemeinsamkeiten) herausgearbeitet und genutzt, um bestehende Sichtweisen zu hinterfragen und das Verständnis für andere Meinungen zu fördern. Damit gegenteilige Standpunkte akzeptiert (statt abgelehnt) werden, ist eine sorgfältige Gestaltung der Prozessschritte erforderlich.

Im Rahmen von Evaluationen kann die dialektische Methode sowohl für die Gewinnung als auch für die Analyse von Daten verwendet werden. Beim *Convergent Interviewing* werden Informationen mit einem Fokus auf die Herausarbeitung signifikanter Unterschiede zwischen den unterschiedlichen Perspektiven eingeholt, wobei die einzelnen Interviews aufeinander Bezug nehmen (zum Prozessablauf siehe Beispiel unten). Bei der *Contradiction Analysis* werden Informationen mit Stakeholdern ausgewertet, um gemeinsam Einsichten auf Basis von ‚Ausreißern‘

⁴ Verwandte Ansätze: „Government and Opposition“ aus den Tiny Tools von NGO-IDEAs arbeitet ähnlich mit zwei Perspektiven plus Beobachtern: <http://www.ngo-ideas.net/narratives/> (07.09.2012). Das Konzept der kollegialen Fallberatung von BeraterInnen (u. a. in Capacity Works, GTZ 2008, 269) fokussiert auf die Selbstreflexion der EvaluatorInnen. Beteiligt sind EvaluatorIn und Externe. Es handelt sich hier um drei Perspektiven, weil die Externen einmal Hypothesen bilden und dann Interventionen planen, also mit zwei verschiedenen Perspektiven auf die Sache blicken.

zu gewinnen. Im Vordergrund steht die Reflexion über Ausnahmen, Widersprüche, Überraschungen und unerklärbare Phänomene. Dieses Verfahren fördert die Artikulation gegenteiliger Standpunkte, um dadurch die Gründe für die Abweichungen und Ausreißer besser zu verstehen.

BEISPIEL:

PROZESSABLAUF VON CONVERGENT INTERVIEWING

Jedes Interview beginnt mit offenen Fragen, damit jeder Interviewpartner zunächst möglichst unbeeinflusst und mit eigenen Worten seine Sichtweise (z. B. zu einem Vorhaben) äußern kann.

Im Anschluss werden sogenannte Sondierungsfragen gestellt, die sich aus früheren Interviews ergeben haben. Damit werden die Antworten anderer Interviewpartner hinterfragt. Die Inhalte der Interviews sind nicht standardisiert, sondern von den Interviewten weitgehend selbst bestimmt.

Abschließend werden die Interviewpartner gebeten, Hauptpunkte zusammenzufassen und weitere Gesprächspartner vorzuschlagen, die andere oder gegensätzliche Standpunkte vertreten.

Der Schlüssel für diese Technik liegt in der Identifizierung von Sondierungsfragen. Dazu müssen die Interviews fortlaufend ausgewertet werden. Während des Interviews wird nach zwei Mustern interveniert: Bei Zustimmung zu früheren Aussagen wird nach Ausnahmen gefragt (ist das immer so?); bei Widerspruch nach Erklärungen (Warum? Unter welchen Rahmenbedingungen?).

Die Interviews innerhalb einer Serie konvergieren im Lauf der Zeit, und zwar in einem doppelten Sinn: Die Struktur wird ähnlich, da wesentlich mehr Zeit für die Sondierungsfragen aufgewendet wird. Der Inhalt wird ebenfalls ähnlich, da sich die Sondierungsfragen in der Regel auf wenige Schlüsselthemen konzentrieren. Insgesamt kann mit dieser Technik sehr effizient befragt werden, die Interviews werden tendenziell kürzer und fokussierter. Es ist allerdings sinnvoll, dass alle Interviews von der gleichen Person durchgeführt werden.

2.3 Systemgrenzen

Systemisches Denken betrachtet zwar Situationen in ihrem Zusammenhang, ist aber keine ganzheitliche Betrachtungsweise, denn es erfordert Unterscheidungen bzw. Abgrenzungen (z. B. zwischen System und Umwelt oder zwischen Gesamtsystem und Teilsystemen). Auch diese Abgrenzungen sind nicht absolut, sondern erfolgen mit Bezug auf einen spezifischen Zweck und aus einer bestimmten Perspektive heraus. Deren Einfluss für die Festlegung von Grenzen soll daher hinterfragt werden, um zu bestimmen, welche Perspektive für die Festlegung von drinnen/ draußen, wichtig/unwichtig etc. ausschlaggebend ist.

Systemische Methoden bieten Vorgehensweisen für das Ziehen von (System-) Grenzen und für eine kritische Reflexion dieser Abgrenzungen an. Sie fördern den bewussten Umgang mit Grenzen, und die Abwägung der Konsequenzen von Abgrenzungen. Damit können auch unterschiedliche Möglichkeiten eröffnet werden, eine Situation als System zu betrachten und somit abzugrenzen.

CRITICAL SYSTEMS HEURISTICS (CSH)

Dieser Ansatz wurde in Großbritannien entwickelt, um die Implikationen von Systemgrenzen besser auszuloten. Jede Abgrenzung hat im Kern eine ethische Dimension, die eng mit Werthaltungen, Interessen und Macht verknüpft ist. Ohne kritische Hinterfragung können diese Dimensionen und ihre Konsequenzen nicht umfassend verstanden werden und es kommt möglicherweise (und unbeabsichtigt) zu Benachteiligungen. CSH ist die umfassendste und eine besonders fundierte Methode, um Systeme passend abzugrenzen sowie die (praktischen und ethischen) Implikationen unterschiedlicher Abgrenzungen zu verstehen.

CSH ist eine Heuristik, d. h. ein analytisches Vorgehen, bei dem mit begrenztem Wissen über ein System mit Hilfe von Mutmaßungen Schlussfolgerungen über das System getroffen werden. Dafür wird eine Reihe von Fragen in einer bestimmten Reihenfolge durchgearbeitet. Dieses Verfahren zur kritischen Reflexion von Systemabgrenzungen beruht auf vier Dimensionen: **(1) Motivation** (Nutznießler, Wertebasis), **(2) Kontrolle** (Machtbasis, Entscheidungsbefugnis), **(3) Wissen** (Relevanz von Expertise) und **(4) Legitimität** (Nachteile und Benachteiligte). Durch die Beantwortung der dazugehörigen Fragestellungen können die für eine Abgrenzung relevanten (aber oftmals verdeckten) Annahmen, Werte und Machtbeziehungen transparent gemacht werden. Damit wird auch auf die Risiken von Vernachlässigungen (z. B. von Stakeholdern, Ressourcen, Expertise) hingewiesen (vgl. ULRICH, REYNOLDS 2010).

BEISPIEL:

CSH IN EINEM PROJEKT ZUR REDUZIERUNG VON HIV/AIDS

Ziel des Projekts war die Reduzierung von HIV/AIDS-Infektionen in einer Bergbauregion. Schwerpunktmäßig wurden lokale Gesundheitsdienste sowie Fortbildung und Mikrokredite zur Schaffung alternativer Einkommensmöglichkeiten für Sexarbeiterinnen unterstützt. Die Umsetzung war mangelhaft und zeigte überraschende, teilweise widersprüchliche Ergebnisse. Die Anwendung von CSH bei der Zwischenevaluation zeigte folgendes auf:

Motivation: Die Nutznießerinnen waren zwar benannt (Sexarbeiterinnen), doch einige der dem Projekt zugrunde liegenden Annahmen entsprachen nicht deren Anliegen. Die Wertebasis des Projekts war daher primär jene der Geldgeber, was auch in den Maßnahmen und Erfolgsindikatoren zum Ausdruck kam.

Kontrolle: Das Projekt hatte insgesamt zu wenig Kontrolle über Ressourcen. Die Verantwortlichkeiten waren schlecht abgegrenzt, so dass einige Schlüsselakteure entweder zu wenige Entscheidungsbefugnisse hatten (z. B. Auszubildende) oder niemandem wirklich verantwortlich waren (im Fall Mikrokredit).

Wissen: Das Projekt beruhte teilweise auf falschen Annahmen über das vorhandene Wissen (z. B. HIV-Wissen von Sexarbeiterinnen, Expertise für Mikrokredit). Informationen über vorhandenes Wissen wurden vorab kaum eingeholt und daher bei der Projektkonzeption nicht berücksichtigt.

Legitimität: Die Interessen wichtiger Stakeholder wurden nicht berücksichtigt (z. B. Klienten der Sexarbeiterinnen, Hotelbesitzer). Dies hatte zur Folge, dass sich diese benachteiligt fühlten und das Projekt nur einen Teil seiner Zielgruppe unter den Sexarbeiterinnen erreichte.

Durch die Anwendung von CSH konnte die Evaluation zwischen Mängeln in der Umsetzung und in der grundsätzlichen Konzeption des Projektes differenzieren.

OUTCOME MAPPING (OM)

Ein wichtiger Aspekt von Systemgrenzen ist die Abgrenzung von Verantwortung. Dabei können die Überlegungen dieses Monitoring- und Evaluations- (M&E) Ansatzes hilfreich sein, der Ende der 1990er Jahre als Reaktion auf die Nachteile herkömmlicher M&E Methoden (z. B. *Logical Framework*) entstanden ist. Entwicklung wird dabei als komplexer Veränderungsprozess aufgefasst, der in erster Linie durch die (Umsetzungs-)Partner bewirkt und durch ein Vorhaben nur unterstützt wird (z. B. durch einen befristeten Zugang zu Ressourcen, Ideen). Diese Sichtweise steht im Einklang mit den Ansprüchen nachhaltiger Entwicklungsarbeit. Bei Planung, Monitoring und Evaluation liegt der Fokus daher auf den Veränderungen (in Verhalten, Beziehungen oder Aktivitäten) von ‚*Boundary Partners*‘, d. h. der Schlüsselakteure, mit denen ein Vorhaben arbeitet oder die es zu beeinflussen sucht. Da bei diesem Ansatz nicht das Erreichen bestimmter Wirkungen, sondern der Beitrag zu Verhaltensänderungen bewertet wird, ist er besonders für Vorhaben geeignet, die primär indirekt wirken (sollen) oder für das Monitoring von Vorhaben zum Aufbau von Institutionen, Kompetenzen und Kapazitäten.

Bei *Outcome Mapping* wird auch explizit anerkannt, dass die verschiedenen Partner mit unterschiedlichen Eigenlogiken und Verantwortlichkeiten arbeiten. Das darin enthaltene Modell der drei Sphären ist gut zur Abgrenzung von Verantwortungsbereichen geeignet: Die beteiligten Akteure werden drei Teilsystemen zugeordnet, je nachdem ob sie für das Vorhaben von Interesse sind, in dessen Einflussbereich liegen oder über das Vorhaben kontrolliert werden können. Dadurch kann auch die Verantwortung von EZ-Akteuren (z. B. Durchführungsorganisationen, Gebern) gegenüber anderen PartnerInnen abgegrenzt werden (vgl. EARL, CARDEN, SMUTYLO 2001)

BEISPIEL: DIE DREI SPHÄREN VON OM IN EINEM „COMMUNITY DEVELOPMENT“ PROGRAMM

Zunächst werden die am Programm Beteiligten nach drei Kategorien unterschieden:

- » **Beneficiaries** (Zielgruppen bzw. Nutznießer des Programms), z. B. die Bewohner eines Landkreises.
- » **Stakeholder** (verfügen über wichtige Ressourcen oder vertreten wichtige Interessen in Bezug auf das Programm).
- » **Boundary Partners** (jene Stakeholder, mit denen das Programm direkt arbeitet oder die beeinflusst werden sollen), z. B. lokale Verwaltungen.

Im Anschluss werden sie in einer Akteurslandkarte (siehe Abb. 5) den drei Sphären zugeordnet. Diese Zuordnung dient u. a. dazu, die Aufgaben der Durchführungsorganisation gegenüber anderen Beteiligten zu präzisieren und Verantwortlichkeiten festzulegen:

- » **Sphere of control:** Die Aufgabenerfüllung der an der Programmdurchführung unmittelbar Beteiligten wird anhand von Zeit-, Ressourcen- und Ergebnisplanung kontrolliert.
- » **Sphere of influence:** Mit den Boundary Partnern werden die von ihnen erwarteten Veränderungen vereinbart (Outcome Challenges) und Indikatoren zu deren Beobachtung festgelegt (Progress markers). Für andere Stakeholder kann das zwar ebenfalls präzisiert, aber nicht vereinbart werden.
- » **Sphere of interest:** Die Anliegen und Interessen weiterer Akteure werden bei der Konzeption berücksichtigt. Das Programm beobachtet die Veränderungen der Akteure während der Umsetzung und modifiziert gegebenenfalls seine Interventionen.

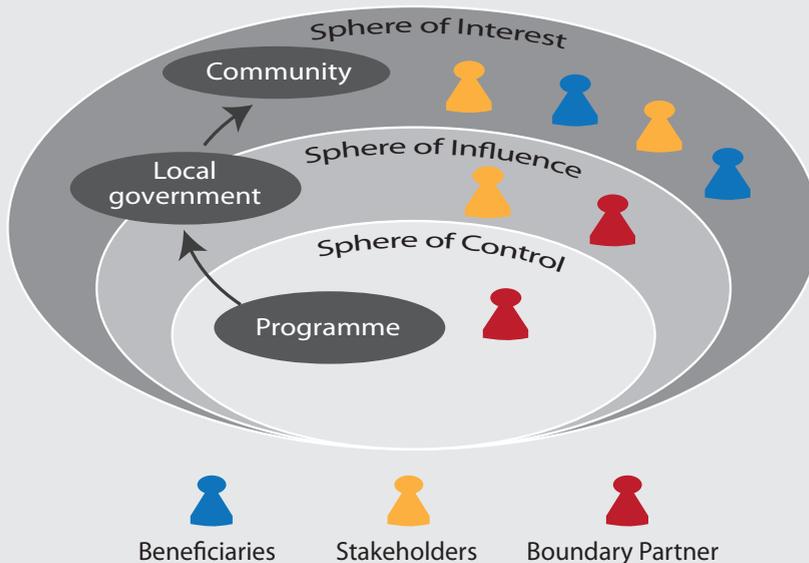


Abb. 5: Outcome Mapping; Quelle: eigene Darstellung

VIABLE SYSTEM MODEL (VSM)

Dieses Modell wendet Erkenntnisse über die Funktionsweise von Organismen auf soziale Systeme an. Ausgehend von der Abgrenzung eines Systems von seiner Umwelt werden die erforderlichen Austauschbeziehungen (von Informationen, Ressourcen etc.) definiert, damit das System seinen Zweck erfüllen kann und in der gegebenen Umwelt langfristig lebens- und entwicklungsfähig bleibt. Die Austauschbeziehungen werden zu Funktionen gebündelt und Akteuren zugeordnet, die wiederum als (ineinander verschachtelte) Teilsysteme voneinander abgegrenzt werden. Deren Verknüpfung ergibt die für eine optimale Leistungsfähigkeit maßgeblichen Beziehungen und Feedback-Schleifen. Dieses Modell eignet sich besonders für die Gestaltung oder Steuerung von Organisationen bzw. Unternehmen in dynamischen Umwelten und ist die konzeptionelle Grundlage für einige systemisch-evolutionäre Managementansätze (z. B. Managementmodell St. Gallen).

Das Modell lebensfähiger Systeme bildet die Mindestanforderungen, für eine nachhaltige Leistungsfähigkeit sozialer Systeme ab. Dazu werden fünf Teilsysteme abgegrenzt, die wiederum drei Managementbereichen zugeordnet werden können (operatives, strategisches und normatives Management). Ferner zeigt das Modell die erforderlichen Beziehungen dieser Teilsysteme an, sowohl untereinander als auch mit ihrer jeweils relevanten Umwelt (vgl. REYNOLDS, HOLWELL 2010; MALIK 1984). Anhand dieses idealtypischen Modells kann eine reale Situation analysiert werden, um daraus Erkenntnisse für Verbesserungen der Leistungs- und Überlebensfähigkeit von Organisationen oder Unternehmen abzuleiten.

BEISPIEL:

STRUKTURIERUNG EINES EZ VORHABENS MITTELS VSM

Die fünf Teilsysteme und ihre Funktionen bzw. Aufgaben (am Beispiel eines EZ Vorhabens):

- 1. Operative Einheiten** erzielen die Outputs und Ergebnisse, steuern sich innerhalb des vorgegebenen Rahmens selbst und beobachten den für sie relevanten Ausschnitt der Umwelt (z. B. Teams für einzelne Maßnahmen/Prioritäten, dezentrale Projektbüros).
- 2. Koordination und Information** unterstützt die Selbststeuerung der operativen Einheiten (z. B. Koordinationsteams, Informationssysteme).
- 3. Management und Kontrolle** achtet auf das Gesamtoptimum und Synergien zwischen den operativen Einheiten, und entscheidet über die Ressourcenverteilung (z. B. Leitungsgremium, Internes Controlling, Monitoring-System).

4. **Zukunftsbezogenes Teilsystem** beobachtet die relevante Umwelt und leitet daraus Anforderungen für das Vorhaben ab (z. B. Lernen und Innovation).

5. **Policy gleicht aus** zwischen gegenwärtigen und zukunftsbezogenen Tätigkeiten sowie zwischen internen und externen Anforderungen und koordiniert die Teilsysteme 3 und 4 (z. B. Vorhabensverantwortlicher).

Anhand dieses idealtypischen Modells wird analysiert, ob und inwieweit das Vorhaben diese Funktionen erfüllt, wer an den einzelnen Teilsystemen beteiligt ist und ob die entsprechenden Informationsflüsse zwischen diesen gewährleistet sind.

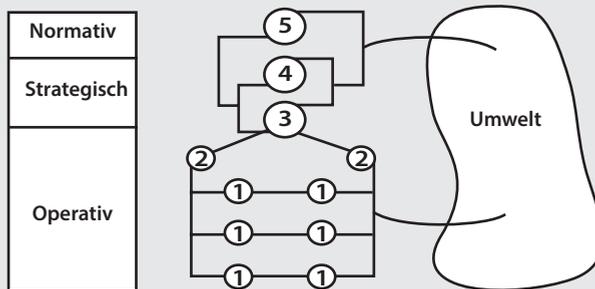


Abb. 6: VSM; Quelle: eigene Darstellung

2.4 Anwendung systemischer Methoden

Die skizzierten Methoden sind ein kleiner Ausschnitt aus der großen Bandbreite der Methoden, die den Konzepten ‚Beziehungen‘, ‚Perspektiven‘ und ‚Grenzen‘ zugeordnet werden können. Die Auswahl geeigneter Methoden sollte sich an der Situation und den jeweiligen Evaluations-Fragestellungen orientieren. Zur besseren Absicherung von Evaluationsergebnissen ist dabei der Grundsatz der Triangulation zu berücksichtigen: systemische Methoden können mit anderen Erhebungs-, Analyse- und Bewertungsmethoden kombiniert und gegengeprüft werden. Dies gilt insbesondere für das Repertoire partizipatorischer Methoden, die vielfach ebenfalls auf systemischen Überlegungen beruhen. Bei einem Multimethodenansatz geht es nicht nur um Methoden in ihrer Gesamtheit; es können durchaus einzelne Elemente verschiedener Methoden (z. B. Prinzipien oder Techniken) in kreativer Weise miteinander verknüpft werden.

Die Anwendung der Systemkonzepte und eine sorgfältige Auswahl der dazu gehörigen Methoden sind allerdings für eine systemische Herangehensweise nicht ausreichend. Es bedarf auch einer spezifischen systemischen Haltung der Beteiligten, insbesondere der EvaluatorInnen.

KENNZEICHEN DER SYSTEMISCHEN HALTUNG

(1) Reflexive Herangehensweise

Handlungen sollten auf Basis vorab gebildeter Hypothesen geplant werden. Die Folgen dieser Handlungen sollten beobachtet und die ursprünglichen Annahmen überprüft werden, um daraus neue Handlungen abzuleiten usw. („systemische Schleife“). Für Evaluationen bedeutet dies, dass sie flexibel genug angelegt sein sollten, um während der Durchführung gegebenenfalls auf neue Erkenntnisse, Entwicklungen oder Anforderungen zu reagieren.

PRAXISTIPP

Während der Planung und Durchführung einer Evaluation kann man eine reflexive Herangehensweise fördern, indem ...

- » *regelmäßiges Feedback mit Auftraggebern stattfindet,*
- » *ein Mitglied des Evaluationsteams explizit den Evaluationsprozess beobachtet und Feedbackrunden im Team stattfinden,*
- » *Spielraum für Adaptierungen in der Evaluation eingeplant wird.*

(2) Zirkuläres Denken

Beim zirkulären Denken handelt es sich um ein Denken in Zusammenhängen, Wechselwirkungen und multiplen Einflüssen (statt linear und monokausal).⁵ Soziale Systeme befinden sich in einer wechselseitigen Abhängigkeit zu ihrer Umwelt und öffnen sich dieser in selektiver Art und Weise, um über diese Verbindungen Einflüsse oder Ressourcen auszutauschen. In Evaluationen sollte daher die Betrachtungseinheit stets aus dem evaluierten System und dem jeweils relevanten Kontext bestehen.

⁵ In der Kybernetik bezeichnet man das Resultat von Rückkopplungsprozessen mit dem Begriff der Zirkularität. Diese beschreibt das zentrale Prinzip kybernetischen Denkens oder kybernetischer Prozesse. Darin wird ein Verhalten einer systemischen Einheit beschrieben, indem die Wirkungen des eigenen Verhaltens rückgekoppelt werden, um das zukünftige Verhalten des Systems direkt und unmittelbar beeinflussen zu können.

PRAXISTIPP

Partizipatorische Planung, Datenerhebung und -analyse fördern zirkuläres Denken, indem sie die Perspektiven eines breiten Spektrums von Stakeholdern berücksichtigen. Bei der Zusammenstellung von Evaluationsteams kann es hilfreich sein, Experten aus dem jeweiligen Projektkontext zu rekrutieren.

(3) Anerkennung bzw. Nutzung der Selbstorganisation

Soziale Systeme operieren relativ autonom; sie sind auf sich selbst bezogen und regulieren sich auf Basis ihrer eigenen Logik und Regeln. Gleichzeitig sind sie von ihrem Kontext abhängig und müssen sich diesem anpassen, um eine stabile Balance zu wahren. Soziale Systeme verändern sich daher fortlaufend. Charakteristisch für diese Selbsterneuerung ist das Auftreten neuer Strukturen, Muster, Eigenschaften im Zeitablauf (Emergenz). Dies bedeutet für Evaluationen, sensibel für Veränderungen (sowohl im evaluierten System als auch in dessen Kontext) zu sein, und zu beobachten, wie das KlientInnensystem auf Impulse bzw. Handlungen der EvaluatorInnen reagiert.

PRAXISTIPP

Auch hier ist es hilfreich, partizipatorische Elemente in den Evaluationsprozess einzufügen und Spielraum für Adaptierungen einzuplanen, um z. B. unerwartete Veränderungen zu beobachten und zu analysieren.

Nach diesem ersten Überblick über die Hauptbestandteile systemischer Ansätze und einiger Anwendungen betrachtet das folgende Kapitel den Evaluationsprozess aus systemischer Sicht. Es umreißt die wesentlichen Implikationen für (1) das Evaluationssystem, (2) das Verständnis von Evaluation als Intervention und als Lernprozess und (3) die Rollen der Evaluierenden.

3 EVALUATION AUS SYSTEMISCHER SICHT

Die Verwendung systemischer Ansätze geht mit einem spezifischen Verständnis von Evaluation einher. Dies wird nachfolgend in Bezug auf vier wesentliche Aspekte skizziert.

3.1 Das Evaluationssystem

Systemisch betrachtet ist Evaluation eine spezielle Form von Beobachtung, die nach professionellen Standards organisiert wird und zumindest zwei beteiligte soziale Systeme umfasst (siehe Abb. 7):

KlientInnen-System: Die Geldgeber und die Durchführenden sowie die Partner eines Vorhabens (d. h. der Evaluationsgegenstand) und weitere Akteure, die an den Evaluationsergebnissen bzw. ihrer Nutzung interessiert sind.

EvaluatorInnen-System: Akteure, die mit der Durchführung der Evaluation beauftragt sind.

Diese Beteiligten konstruieren gemeinsam das Evaluationssystem, das für die Dauer der Evaluation eingerichtet und mittels Vertrag festgelegt wird. EvaluatorInnen vereinbaren mit den KlientInnen ihre für die jeweilige Evaluation geltenden Verantwortlichkeiten, sowie die Bedingungen und Kriterien für die Interaktion mit anderen *Stakeholdern* (z. B. deren Beteiligung an der Erhebung oder Interpretation von Daten). Hier kann es – auch innerhalb der Anwendung systemischer Ansätze – erhebliche Unterschiede geben, abhängig z. B. vom Auftraggeber, der Evaluationskultur im Einsatzgebiet oder den von bestimmten Beteiligten gemachten Erfahrungen.

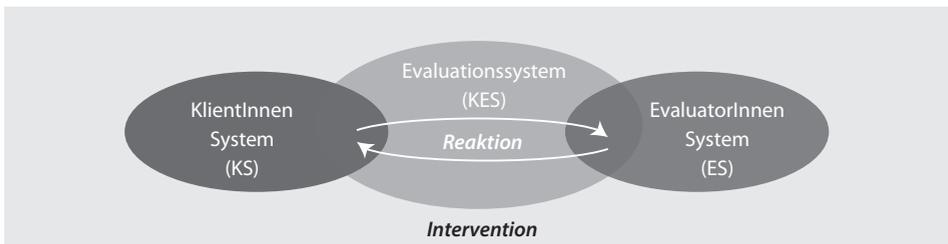


Abb. 7: Evaluationssystem; Quelle: eigene Darstellung

In der Entwicklungszusammenarbeit (EZ) werden Evaluationen häufig nicht vom Vorhaben, sondern von den Geldgebern beauftragt, womit in der Durchführung ein spezifisches Dreiecksverhältnis mit getrennten Interessen und Verantwortlichkeiten entsteht. In derartigen Fällen empfiehlt es sich, das KlientInnen-System weiter aufzuteilen und zwischen Auftraggebersystem und evaluiertem System zu differenzieren.

3.2 Evaluation als Intervention

Jede Evaluation ist eine Intervention in bestehende soziale Systeme, die sowohl anregend als auch (ver)störend wirken kann. Art und Ausmaß der Intervention sind abhängig vom Adressaten der Intervention und nicht vollständig vorhersehbar. Bereits eine einzige Frage in einem Interview kann daher erhebliche Konsequenzen auslösen. Jede Intervention kann wiederum zu Reaktionen (positiv oder negativ) führen, die auf den Intervenierenden zurückwirken.

PRAXISTIPP

EvaluatorInnen sollten sich des Interventionscharakters von Evaluationen bewusst sein und Risiken und Chancen einer Beeinflussung des Systems bei der Evaluationsvorbereitung – möglichst unter Einbeziehung unterschiedlicher Beziehungen – identifizieren. Auch im Rahmen ethischen Handelns sollten Evaluationsdesign und Methodenauswahl Risiken erkennen und minimieren.

Jedes System ist zugleich Umwelt für alle anderen Systeme; dies gilt auch für die an einer Evaluation beteiligten sozialen Systeme. Was immer im Evaluationssystem passiert, kann Auswirkungen auf die konstituierenden Subsysteme haben und umgekehrt. Sie stehen in einem zirkulären Zusammenhang und sind wechselseitig voneinander abhängig. EvaluatorInnen agieren z. B. in einem Rahmen, der häufig vom Auftraggeber-System vorgegeben ist. Ihre Handlungsfähigkeit wird von der Bereitschaft bzw. Fähigkeit der verschiedenen Stakeholder zur Mitwirkung beeinflusst (was wiederum Konsequenzen für die Anwendbarkeit systemischer Methoden haben kann). Das Wissen um diese Abhängigkeiten erlaubt es EvaluatorInnen und Auftraggebern, Risiken von Parteilichkeit zu reduzieren.

3.3 Evaluation als Lern- und Reflexionsprozess

Für einen gemeinsamen Lern- und Reflexionsprozess der an einem Vorhaben Beteiligten ist eine Kombination von interner Reflexion und externen Impulsen durch die EvaluatorInnen zweckmäßig. Dabei geht es nicht nur um die kurzfristige Bewältigung isolierter Problemstellungen (Anpassungslernen), sondern auch um Erkenntnisse für künftiges erfolgreiches Handeln. Dies schließt die Infragestellung von Werten und Handlungsmustern sowie den konstruktiven Umgang mit Unterschieden und Konflikten ein (Veränderungslernen).

Die konstruktive Auseinandersetzung mit Unterschieden im evaluierten System ermöglicht einen Lernprozess, der die Problemlösungskapazität des Systems erhöht. Um die für die Erzielung von Lerneffekten erforderlichen Informationen zu liefern, sollte Evaluation konsequent anwendungsorientiert sein. Die beabsichtigten Anwendungen sind bereits zu Beginn zu identifizieren; NutzerInnen der Ergebnisse sind von Anfang an an der Evaluation zu beteiligen.

PRAXISTIPP

Bei der Vorbereitung einer Evaluation empfiehlt es sich zu klären, wer die Evaluation und ihre Ergebnisse nutzen wird. Die Lernbedarfe der NutzerInnen und die Möglichkeiten, diese im Rahmen der Evaluation zu erfüllen, sollten im Rahmen der Evaluationsvorbereitung besprochen werden, z. B. in Einzelinterviews oder in einem Inception-Workshop, der die verschiedenen Perspektiven und Anliegen unterschiedlicher Beteiligter in Bezug auf die Evaluation verdeutlicht.

3.4 Die Rolle von EvaluatorInnen

Da Evaluation eine organisierte Form der Beobachtung mit verschiedenen Beteiligten darstellt, sind die Erkenntnisse und Ergebnisse beobachterabhängig. Dies gilt auch für EvaluatorInnen, denn diese nehmen eine bestimmte Sichtweise ein und können nach systemischer Auffassung nicht den Anspruch erheben, völlig unabhängig und objektiv Bewertende zu sein – schlichtweg weil es keine Objektivität geben kann. In Fällen, wo Auftraggeber eine **objektive** Sichtweise wünschen, sollte geklärt werden, was damit gemeint ist. Oft wird vor allem **Neutralität** erwartet, d. h. ein möglichst unparteiischer und ausgewogener Blick.

EvaluatorsInnen sollten bei ihrer subjektiven Synthese die Sichtweisen der verschiedenen Akteure berücksichtigen und zumindest signifikante Unterschiede erkennbar machen. Sie sollten ihre Subjektivität reflektieren und, wo relevant, transparent machen. Um trotz der bestehenden Abhängigkeiten die Einhaltung professioneller Standards sicherzustellen, muss insbesondere die Kompetenz der EvaluatorsInnen zur Gestaltung und Reflexion des methodischen Vorgehens sowie zur Darstellung der Ergebnisse gewährleistet sein (und gegenüber den anderen Beteiligten auch durchsetzbar).

Die Anwendung systemischer Ansätze stellt daher spezifische Anforderungen an EvaluatorsInnen. Sie sollten neben der generellen Fachkompetenz für Datensammlung und -analyse auch über Kenntnisse der systemischen Beratung und Methoden verfügen. Sie sollten in der Lage sein, **unterschiedliche Rollen** einzunehmen (z. B. Moderator, Coach, Impulsgeber) und diese während der Durchführung zu reflektieren. Allerdings liegt die Ausübung dieser Rollen nicht alleine im Ermessen der EvaluatorsInnen, sondern sollte mit den anderen Beteiligten (insbesondere dem Auftraggeber) explizit vereinbart werden. Nicht zuletzt müssen EvaluatorsInnen auch in der Lage sein, die Evaluationsprozesse so zu gestalten, dass die Evaluationsziele erreicht werden können – und sie sollten rechtzeitig erkennen, wenn dies gefährdet ist.

Im Anschluss an diese Betrachtungen zum Evaluationsprozess aus systemischer Sicht und seinen Implikationen stellen wir im folgenden Kapitel systemische Herangehensweisen zur Analyse von Wirkungen und Kausalität als zentrales Element von Evaluationen vor.

4 WIRKUNGEN UND KAUSALITÄT AUS SYSTEMISCHER SICHT

Die Analyse von Wirkungen ist zentraler Bestandteil vieler Evaluationen. Innerhalb der Wirkungsanalyse ist die Analyse der Kausalitäten zumeist die schwierigste Aufgabe. Systemische Ansätze bieten Zugänge und Verfahren, die sich für die Anwendung auf komplexe Situationen eignen in denen es keine offensichtlichen Ursache-Wirkungs-Muster gibt. Daher wird diesen Aspekten hier besonderes Augenmerk geschenkt.

4.1 Systemische Betrachtung von Wirkungen

Aus systemischer Sicht kann grundsätzlich zwischen zwei Arten von Wirkungsbetrachtung unterschieden werden:

WIRKUNG **VON** SYSTEMEN (SYSTEMANALYSEN)

- » Leitfragen: Wie wirken einzelne Systemelemente zusammen? Wie wirken Systeme nach außen?
- » Wichtige Analysedimensionen: Subsysteme und ihre Beziehungen; Beiträge der Subsysteme zum Ganzen (z. B. Oberziele); Strukturen, Prozesse und Regeln.
- » Ziel ist die Systemoptimierung (z. B. Kooperationssystem zur Erreichung von Wirkungen).

Die Optimierung oder Lösung wird durch (logische, mathematische) Modelle ermittelt. Dies erfordert definierte Ziele, darstellbare Situationen sowie eine eindeutige Abgrenzung der Subsysteme und Problemsituationen; dafür stehen Methoden aus Systems Engineering oder Operational Research zur Verfügung (vgl. HABERFELLNER et al. 1999).

WIRKUNG **AUF** SYSTEME (SYSTEMISCHE WIRKUNGSANALYSE)

- » Leitfrage: Wie wirken Interventionen (z. B. Vorhaben) auf soziale Systeme?
- » Die Wirkungserfassung basiert auf einigen Grundannahmen über das Verhalten sozialer Systeme: wechselseitige Abhängigkeit von System und Kontext, operative Geschlossenheit (selbstbezüglich, autonom), selektive Öffnung (Verbindung zu Umwelt/Kontext), ‚Emergenz‘, d. h. das Auftreten neuer Strukturen, Muster, Eigenschaften im Zeitablauf. Während mechanische Systeme (z. B. Maschinen) trivial im Sinne von linear/voraussagbar funktionieren, wirken Interventionen in sozialen Systemen auf eine nicht-triviale Art und Weise (vgl. VON FOERSTER 1999).
- » Ziel ist die Beobachtung und Optimierung der Wirkung von Interventionen.

Da soziale Systeme Interventionen gemäß ihrer Eigenlogik verarbeiten, sind der interne Zustand des Zielsystems (in das interveniert wird) und die jeweilige Beziehung zum Kontext ausschlaggebend für die Wirkung von Interventionen. Dies erfordert eine dynamische, kontextbezogene Betrachtung von Wirkungen. Die Analyse der Wirkungen erlaubt zudem Rückschlüsse auf das soziale Zielsystem (vgl. Abb. 8).

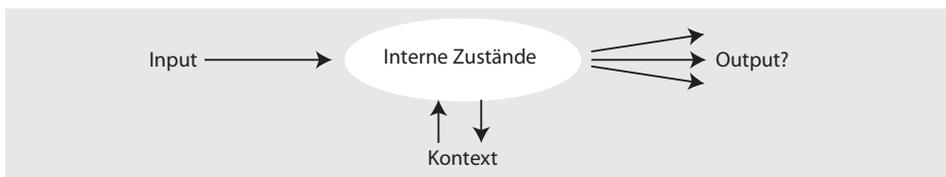


Abb. 8: Nicht-triviale Wirkungsweisen⁶; Quelle: VON FOERSTER 1999

Eine derartige Betrachtungsweise von Wirkungen hat eine Reihe von praktischen Konsequenzen für Evaluationen:

DYNAMISCHE BETRACHTUNG VON EFFEKTEN

Das Eintreten von Effekten soll nicht auf Grundlage eines statischen Soll-Ist Vergleichs erfolgen, sondern deren Dynamik im Zeitablauf berücksichtigen. In der Regel erweist sich nur ein Teil der geplanten Effekte als tatsächlich plan- und durchführbar. Andererseits kommen neue Effekte hinzu (positiv und/oder negativ), die bei der Planung nicht mitgedacht waren. Beides zusammen ergibt die tatsächlich realisierten Effekte.

⁶ Die Bezeichnungen „Input“ und „Output“ in diesem Schaubild sind die Begriffe wie sie VON FOERSTER verwendet; sie entsprechen nicht der bei theoriebasierten Evaluationen verwendeten Terminologie.

DIFFERENZIERTER UMGANG MIT ABWEICHUNGEN

Abweichungen sollten nicht von vorneherein als negativ angesehen werden („Korrektur-Reflex“), sondern ihre Gründe und Motive zunächst eingehender untersucht werden. Abweichungen haben oftmals eine wichtige Funktion im Entwicklungsprozess, denn sie können Schwächen, bisher unerkannte Potentiale und mögliche Alternativen aufzeigen. Damit können Unterschiede zwischen Soll und Ist, Plan und Umsetzung wichtige Erkenntnisse und Hinweise für sinnvolle Lösungen liefern.

BEACHTUNG VON ÜBERRASCHENDEN BEFUNDEN

Falls die Datenanalyse signifikante Ausnahmen, Diskontinuitäten, Widersprüche oder Rätsel ergibt, sollten diese einer eingehenden Betrachtung unterzogen werden. Um diese Ergebnisse nicht als isolierte Phänomene zu behandeln, sollten sie der Handlungslogik des Vorhabens zugeordnet werden. Ähnlich wie die Abweichungen können sie Hinweise auf das Entstehen und die Bewältigung von Veränderungen bzw. neuen Herausforderungen geben und Anhaltspunkte für die Anpassung liefern.

IDENTIFIZIERUNG VON NICHT INTENDIERTEN WIRKUNGEN/PHÄNOMENEN

Die Betrachtung einer Situation aus unterschiedlichen Perspektiven erhöht die Wahrscheinlichkeit, dass ungeplante Effekte oder nicht-lineare Entwicklungen entdeckt werden. Auch unterschiedliche Systemabgrenzungen, insbesondere eine Erweiterung des Betrachtungsfelds, sind hilfreich.

LÖSUNGSSYSTEM NICHT IDENTISCH MIT PROBLEMSYSTEM

Auch Lösungen können als Systeme aufgefasst und dargestellt werden. Diese sind nicht deckungsgleich mit dem Problemsystem und zumeist kleiner. Sie erfordern häufig andere bzw. neue Elemente oder Beziehungen. Eine gemeinsame Betrachtung dieser beiden Systeme ist daher nicht zwingend erforderlich. Bei der lösungsorientierten Vorgehensweise wird z. B. auf eine genauere Problemanalyse völlig verzichtet und direkt auf die Lösungsfindung fokussiert. Der Solution Focus (vgl. JACKSON, MCKERGOW 2002) bewährt sich bei Evaluationen unter anderem, wenn die Beteiligten in eine Problemsituation emotional sehr verstrickt sind oder die Ursachen für aktuelle Probleme weit zurückliegen.

4.2 Systemische Wirkungsanalysen

Diese bestehen im Wesentlichen aus drei aufeinanderfolgenden Prozessschritten:

ABBILDUNG/VISUALISIERUNG VON SACHVERHALTEN

Die Erfassung der für die jeweilige Situation relevanten Elemente und Beziehungen ergibt einen ersten Überblick. Die Darstellung sollte möglichst umfassend, anschaulich und auch für Außenstehende leicht erfassbar sein. Sie wird relativ rasch (z. B. in Form von *Brainstorming*) erstellt und bildet das Ausgangsmaterial für die weiteren Überlegungen. Geeignete Methoden sind z. B. *Rich Pictures*, *Mind Mapping*, *Cognitive Mapping*, *Concept Mapping*.

MODELLIERUNG (ALS SYSTEM)

Die Situation wird mit Bezug auf einen Untersuchungszweck bzw. eine Fragestellung strukturiert. Dazu werden die vorhandenen Abbildungen neu konfiguriert, d. h. besonders relevante Elemente bzw. Faktoren ausgewählt und die Komplexität der Systemdarstellung reduziert. Dies geschieht mithilfe von Diagrammen (z. B. Einfluss-, Wirkungs-, Fluss- oder Netzwerkdiagramme) mit unterschiedlichen Funktionen und Aussagen. Es können auch umfassendere Modellierungsmethoden verwendet werden, z. B. das *Viable System Model* (Schwerpunkt auf die Beziehung zwischen Teilsystemen) oder die *Soft Systems Methodology* (zur Modellierung eines Systems aus unterschiedlichen Perspektiven).

ANALYSE/INTERPRETATION DES MODELLS

Dabei werden die Stärken und Schwächen bzw. die Potentiale einer Situation näher untersucht. Auf der Grundlage des Modells können insbesondere relevante Faktoren oder Ansatzpunkte für Veränderungen identifiziert und kreativ über Lösungen nachgedacht werden (z. B. bei verzwickten oder kontroversen Sachverhalten). Die dabei verwendeten Methoden hängen eng mit der Art der Modellierung zusammen: Einflussmatrix bei Einfluss- oder Wirkungsdiagrammen, (Verhaltens-) Simulation bei Flussdiagrammen, Kennzahlenanalyse bei Netzwerkdiagrammen, Analyse von Beziehungen beim *Viable System Model*.

BEISPIEL EINER SYSTEMISCHEN WIRKUNGSANALYSE

Zur Evaluation von Wirkungen beruflicher Bildung in Ghana wurde ein systemisches Verfahren entwickelt, das folgende Schritte beinhaltet: **1) Systembeschreibung, 2) Problem- und Bedarfsanalyse, 3) Formulierung von Annahmen, Zielen, Wirkungshypothesen, 4) Strategieentwicklung, 5) Implementierung, Monitoring, 6) Evaluation.** Bei der Systembeschreibung wird zunächst ein Wirkungsgefüge entworfen (Definition der Systemgrenzen; Bestimmung zentraler Systemelemente und ggf. von Subsystemen).

Die Beziehungen zwischen Elementen in und zwischen den Systemen wurden mithilfe einer Wechselwirkungsmatrix (auch Einflussmatrix nach VESTER 2005) bestimmt (siehe Tab. 2). Die Wechselwirkungen der einzelnen Variablen wurden im Rahmen von Gruppendiskussionen von Experten geschätzt. Es zeigte sich, dass die Zahl der Variablen auf maximal 15 beschränkt sein sollte, um den zeitlichen Rahmen nicht zu sprengen. Die Wirkungen werden auf einer fünfteiligen Skala angeordnet, im Beispiel ist 0,1 der niedrigste und 2,0 der höchste Wert. Die Einflussstärke auf ein Element wird in der Passivsumme dargestellt (zeigt, wie empfindlich die Variable auf Veränderungen reagiert); die Einflussstärke, die von einem Element ausgeht, in der Aktivsumme (zeigt, wie stark die Variable andere Elemente bzw. das System beeinflusst). Nach der Prüfung der wechselseitigen Einflussstärken aller Elemente werden diese auf den Grad ihrer Vernetzung im Gesamtsystem geprüft. Hierbei wird nach aktiven (hohe Aktiv- und niedrige Passivsumme), reaktiven (niedrige Aktiv-, hohe Passivsumme), kritischen (hohe Aktiv-, Passivsumme), puffernden und neutralen Elementen (niedrige Aktiv-, Passivsumme) unterschieden.

Mithilfe des Instruments können Steuerungshebel (Elemente mit hoher aktiver Vernetzung) im System identifiziert sowie solche Einflussgrößen aufgespürt werden, die mit ihrer Trägheit zur Stabilisierung des Systems in einem kritischen Veränderungsprozess beitragen (Elemente mit hoher passiver Vernetzung). So wurde z. B. der Beschäftigungserfolg der Absolventen beruflicher Bildung (orangefarbene Zeile) als kritische Variable mit besonders starker Gesamtvernetzung (hohe Aktiv- und Passivsumme) identifiziert. Steuerungshebel und stabilisierende Elemente werden bei der Hypothesenentwicklung besonders berücksichtigt und zum Zweck der Wirkungsevaluation auf der Mikroebene modelliert und operationalisiert. Der Schwerpunkt des Einsatzes des Instruments liegt in diesem Falle im Prozess der Zielformulierung und Hypothesenentwicklung.

System	No.	Influence Factor	1	2	3	4	5	6	7	8
TVET-system	1	Vision, training approach		1,5	1,5	1,5	1,0	1,5	0,1	0,1
	2	Organisational structure	1,0		2,0	1,5	2,0	2,0	1,0	0,5
	3	Staff / Ownership	1,0	2,0		1,5	1,0	2,0	0,1	0,1
	4	Technology	1,0	1,0	1,0		1,0	2,0	0,1	0,1
	5	Ressources	1,0	1,5	1,5	1,5		2,0	0,1	0,1
	6	Quality of lessons	1,0	1,0	1,5	1,0	1,0		0,1	0,1
Labour market demand (enter-prises)	7	Size, sector of enterprises	1,5	1,0	1,0	1,0	1,0	1,0		2,0
	8	Required skills, qualifications	1,5	1,5	1,5	2,0	1,0	2,0	1,0	
	9	Technology	1,0	1,0	1,0	2,0	1,0	1,0	1,5	2,0
Labour market supply (graduates)	10	Prises/salaries	0,5	0,5	0,5	0,5	0,5	0,5	2,0	1,5
	11	Acces to markets graduates	0,5	0,5	0,5	0,5	0,5	1,0	2,0	2,0
	12	Associations of graduates	0,5	0,5	0,5	0,5	0,1	0,1	1,0	1,0
	13	Employment succes of graduates target groups	0,5	0,5	0,5	0,5	0,5	1,0	1,5	1,0
Educational system	14	Quality primary, secondary education	0,1	0,1	1,0	1,0	0,1	1,5	1,0	1,0
	15	Quality teacher training	0,5	0,5	2,0	1,0	0,5	2,0	1,0	1,0
Population	16	Movement	0,5	0,5	0,5	0,5	0,5	0,5	1,5	1,0
Other social elements	17	educational, economic background target groups, their families	0,1	0,1	0,5	0,1	0,1	2,0	0,1	0,1
	18	Abilities, personal contatcs target groups	0,1	0,1	0,5	0,1	0,1	1,0	0,1	0,1
Political system	19	Cooperation national, regional, district ass.	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
	20	Cooperation civic society	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
	21	Cooperation traditional leaders	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5
Culture and values	22	Status of the profession "teacher"	0,5	1,0	1,0	1,0	1,0	2,0	0,5	0,5
	23	Acceptance of TVET	1,0	0,5	1,0	1,0	1,0	2,0	0,5	0,5
Passive sum			16,3	17,7	22,0	21,2	16,4	26,1	17,7	17,2
Activity ratio (AS/PS)			1,19	1,39	0,9	0,91	1,21	0,74	1,42	1,34
Determination (rating) of the relationships between all the elements										
2,0 = very strong influence 1,5 = strong influence 1,0 = moderate influence										

9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Active sum	Level of interrelation (AS*PS)
1,0	0,1	0,5	0,5	1,5	0,1	1,0	1,0	1,0	1,5	1,0	1,0	0,5	0,5	1,0	19,4	316,2
1,0	1,0	1,0	0,5	1,5	0,1	1,0	1,0	1,0	1,5	1,0	1,0	0,5	1,5	1,0	24,6	435,4
1,5	0,5	0,5	0,5	1,5	0,1	1,0	1,0	1,0	1,5	0,5	0,5	0,5	0,5	1,0	19,8	435,6
2,0	1,0	1,0	0,5	1,5	0,1	1,0	0,5	1,0	1,5	0,5	0,5	0,5	0,5	1,0	19,3	409,2
1,0	0,1	0,5	0,5	1,5	0,1	1,0	1,0	1,0	1,5	0,5	0,5	0,5	1,5	1,0	19,9	326,4
1,5	0,5	1,0	0,5	2,0	0,1	1,0	1,0	1,0	2,0	0,5	0,5	0,5	0,5	1,0	19,3	503,7
2,0	2,0	2,0	0,5	2,0	0,1	0,1	1,5	1,5	0,5	1,0	1,0	1,0	0,5	1,0	25,2	446,0
2,0	1,5	1,5	0,5	2,0	1,0	1,0			0,5	0,5	0,5	0,5	0,5	0,5	23,0	395,6
	2,0	2,0	0,5	1,5	0,5	1,5	1,0	1,0	0,5	0,5	0,5	0,5	0,5	1,0	24,0	568,8
2,0		2,0	1,0	2,0	0,1	0,1	1,5	1,5	0,5	1,0	1,0	1,0	0,1	0,5	20,8	432,6
1,5	2,0		1,0	2,0	0,1	0,1	1,5	2,0	1,0	1,0	1,0	1,0	0,1	0,1	21,9	538,7
0,5	1,0	1,0		1,0	0,1	0,1	0,5	1,0	1,0	1,0	1,0	0,5	0,5	0,5	13,9	190,4
1,0	1,0	1,0	1,5		0,1	0,5	1,5	2,0	1,5	1,0	1,0	1,0	0,5	0,1	19,7	669,8
1,0	1,0	1,0	1,0	1,5		2,0	1,0	1,5	1,5	1,0	1,0	1,0	1,0	1,0	20,7	219,4
1,0	1,0	1,0	0,5	1,5			1,0	1,5	1,5	1,0	1,0	1,0	1,0	1,5	21,0	388,5
1,0	1,5	1,5	0,5	1,5	1,0	0,1		1,5		1,0	1,0	1,0	0,5	0,5	18,1	407,3
0,1	1,0	1,5	0,1	2,0	0,5	0,5	2,0		2,0	0,5	0,5	1,0	1,0	1,0	16,8	436,8
0,1	1,0	2,0	0,1	2,0	1,0	0,5	2,0	2,0		0,5	0,5	0,5	0,1	1,0	15,4	400,4
1,0	1,0	1,0	1,0	1,5	2,0	2,0	1,0	1,0	1,5		1,5	1,5	1,0	1,0	26,0	481,0
1,0	0,5	1,0	1,0	1,0	0,5	1,0	1,0	1,0	1,5	1,5		1,5	1,0	1,0	21,5	397,8
0,5	0,5	1,0	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,5	1,5		1,0	1,0	15,5	271,3
0,5	0,1	0,1	0,5	1,0	2,0	2,0	0,5	1,0	1,0	1,0	1,0	0,5		1,0	19,7	311,3
0,5	0,5	0,5	0,5	1,5	0,5	0,5	0,5	0,5	1,0	0,5	0,5	1,0	1,5		17,5	327,3
23,7	20,8	24,6	13,7	34,0	10,6	18,5	22,5	26,0	26,0	18,5	18,5	17,5	15,8	18,7		
1,01	1,00	0,89	1,01	0,58	1,95	1,14	0,80	0,65	0,59	1,41	1,16	0,89	1,25	0,94		
0,5 = weak influence																0,1 = very weak influence

Tab. 2: Einflussmatrix – Partizipative Systemanalyse; Quelle: Horn (2011)

4.3 Systemische Ansätze zur Kausalanalyse

Die Leitfrage für empirisch-wissenschaftliche Kausalanalysen lautet: „Wie und in welchem Ausmaß lassen sich die Wirkungen auf die Intervention zurückführen?“ Dazu wird versucht, die Ursachenfaktoren für Wirkungen möglichst eindeutig zu bestimmen. Sie werden als Beziehungen zwischen dem Vorhaben (unabhängige Variable) und den erfassten Wirkungen (abhängige Variable) definiert. Der kausale Zusammenhang zwischen diesen Variablen wird möglichst exakt gemessen. Ferner wird eine Wirkungsbilanz erstellt, wobei Nettowirkungen eines Vorhabens durch die Berücksichtigung von externen Faktoren und Designeffekten ermittelt werden (vgl. STOCKMANN 2006).

Das Verfahren besteht in einem Vergleich zwischen der von einem Vorhaben beeinflussten **Versuchsgruppe** und davon unabhängigen **Kontroll- oder Vergleichsgruppen**. Die Wirkung wird anhand der Veränderungen bei den abhängigen Variablen festgestellt und dem Vorhaben ursächlich zugeschrieben (*attribution*). Der Fokus liegt auf der internen Gültigkeit (Validität); diese wird durch den Ausschluss von Störvariablen oder Designeffekten sichergestellt. Dieser Ansatz, der aus den Naturwissenschaften abgeleitet ist, setzt randomisierte Kontrollen, konstruierte Kontrollen (*matching*) und statistische Kontrollen ein.

Bei der Erfassung der Wirkung auf Systeme sind derartige Kausalanalysen nur begrenzt anwendbar, denn die erforderlichen eindeutigen Kausalitäten bzw. Beziehungen zwischen Variablen sind selten oder schwierig zu erfassen. Gleiches gilt für die zur Berechnung der Nettowirkungen eines Vorhabens erforderliche Identifizierung der Effekte anderer Faktoren. Weiterhin besteht die Gefahr, vernetzte Wirkungsgefüge in unangemessener Weise zu reduzieren bzw. eingetretene Wirkungen über Gebühr einzelnen Faktoren oder AkteurInnen zuzurechnen. Nicht zuletzt haben derartige Analysen begrenzte Aussagekraft über die ‚externe Gültigkeit‘, d. h. die Übertragbarkeit auf andere Situationen.

BERÜCKSICHTIGUNG VON WIRKUNGSWEISEN UND KAUSALITÄTSMUSTERN

Bei der Auswahl geeigneter Verfahren sollten vor allem die Wirkungsweise des Vorhabens und die vorherrschenden Kausalitätsmuster berücksichtigt werden. In Bezug auf die Wirkungsweise kann die in der Einleitung dieses Papiers erwähnte Klassifizierung von Komplexität herangezogen werden. Tabelle 3 enthält eine Übersicht der wichtigsten Aspekte von Wirkungsweisen in den drei Domänen, die dabei von Relevanz sind.

Die empirisch-wissenschaftliche Kausalanalyse ist besonders für Situationen geeignet, die der Domäne des ‚Einfachen‘ zuzuordnen sind, z. B. bei einer kurzfristigen Erhöhung von Ernteerträgen oder der Verhinderung somatischer Krankheiten durch einzelne Maßnahmen (z. B. Einsatz von Düngemitteln bzw. Impfstoffen). Die Domänen ‚kompliziert‘ und ‚komplex‘ haben aber andere Kausalitätsmuster und erfordern daher auch andere Ansätze zur Kausal- und Wirkungsanalyse.

	Einfach	Kompliziert	Komplex
Notwendig (zur Erzielung der Wirkung)	Die einzige Weise, Wirkung zu erzielen	Eine von mehreren Möglichkeiten, die beabsichtigten Wirkungen zu erzielen	
		Im Voraus planbar	Nur im Nachhinein erkennbar (evtl. im Ansatz vorhersagbar)
Hinreichend (zur Erzielung der Wirkung)	Reicht aus, um die Ziele zu erreichen; wirkt bei allen AkteurInnen	Erzielt seine Wirkungen nur in Kombination mit anderen Interventionen, nur für bestimmte Akteure oder unter spezifischen Rahmenbedingungen	
		Im Voraus planbar	Nur im Nachhinein erkennbar (evtl. im Ansatz vorhersagbar)
Wirkungsverlauf	Konstantes, lineares Verhältnis zwischen Aufwand und Ergebnis	Veränderungsverlauf ist vorhersagbar, aber erfordert Fachwissen	Veränderungsverlauf ist nicht vorhersagbar und verläuft unstetig
Nicht beabsichtigte Wirkung	Können vorhergesagt und beobachtet werden	Unter bestimmten Bedingungen vorhersehbar (erfordert Expertise)	Nicht vorhersehbar, Identifizierung ist nur im Nachhinein möglich

Tab. 3: Aspekte von Wirkungsweisen; Quelle: nach FUNNELL, ROGERS (2011)

Eine aktuelle Studie des englischen *Department for International Development* unterscheidet diesbezüglich zwischen zwei Kausalitätsmustern (DFID 2012):

ZUSAMMENWIRKEN MEHRERER FAKTOREN

Das Vorhaben wird hierbei nicht isoliert betrachtet, sondern wirkt nur zusammen mit anderen Faktoren.⁷ Dabei wird einerseits untersucht, ob und in welchem Ausmaß das Vorhaben für die Erzielung einer Wirkung notwendig ist, d. h. welche

⁷ Dies entspricht dem Konzept des *Dynamic Fit*, mit dem Wirkungen aus dem wechselseitigen Zusammenspiel unterschiedlicher Variablen abgeleitet werden (vgl. BAGADION 1999).

Elemente zu welchem Zeitpunkt in welchem Umfang und in welcher Art verfügbar waren, und welchen Beitrag sie zur Wirkung des Programms geleistet haben. Andererseits wird untersucht, ob dies für die Wirkung hinreichend ist oder ob dazu weitere Faktoren erforderlich sind. Dabei kann unterschieden werden nach Faktoren, die den Boden bereiten, die auslösend für eine Wirkung sind oder diese behindern: Impfungen z. B. sind ein auslösender Faktor, denn sie reduzieren unmittelbar die Ansteckungsraten von Krankheiten. Die Bereitstellung von Schulbüchern hingegen erfordert andere Faktoren, die den Boden für eine gute Ausbildung bereiten (z. B. regelmäßiger Schulbesuch, Motivation zum Lernen, gut erreichbare Schulen oder gute Lehrer). Deren jeweilige Kombination ist für die Qualität der Ausbildung ausschlaggebend – und einige dieser Faktoren (z. B. Schulen, Lehrer) sind besonders kritisch, denn ohne sie bliebe beispielsweise die Bereitstellung von Schulbüchern wirkungslos.

Von besonderem Interesse sind dabei nicht-lineare multiple Kausalitäten (nach dem englischen Kürzel INUS⁸-Bedingungen genannt): Hier wird die Wirkung durch eine Kombination von Faktoren erreicht (*causal package*), die jeweils individuell notwendig und im Verbund für die Wirkung hinreichend sind. Feuer kann z. B. durch Reiben eines Zündholzes auf Phosphor entstehen oder durch eine brennende Zigarette und Benzindämpfe. Obwohl für die Entstehung von Feuer notwendig, reicht keiner dieser Faktoren (Zündholz, Phosphor, Zigarette, Benzindämpfe) alleine dafür aus – aber sie sind jeweils Teil einer ausreichenden Kombination.

GENERATIVE KAUSALITÄT

Die Analyse von Bedingungen bzw. Faktoren sagt noch nichts über die Art ihres Zusammenwirkens, z. B. deren Reihenfolge oder Beziehung aus. Dazu müssen die Bedingungen zu Konfigurationen zusammengefasst werden, die das Eintreten einer Wirkung erklären. Es geht nicht um einzelne Bedingungen (im Sinne von Variablen), sondern um den gesamten Verlauf von der Intervention zur Wirkung. Dies entspricht der Programmtheorie bei theoriebasierten (Evaluations)Ansätzen. Zur Darstellung können Wirkungsketten oder -gefüge verwendet werden.

In sozialen Systemen werden diese Kombinationen aktiv herbeigeführt („generiert“) und die einzelnen Bedingungen „manipuliert“. Die Wirkung hängt daher auch von den Handlungen und Entscheidungen der beteiligten Akteure ab. Ein Rezept und die dazugehörigen Zutaten reichen nicht aus; jemand muss sie zubereiten, um eine Speise zu erhalten – und dies erklärt auch unterschiedliche Ergebnisse bei den Speisen. Das jeweilige soziale System, z. B. die an der Durchführung eines Vorhabens beteiligten Akteure, beeinflusst dessen Wirkungsweise. Die Gesamtheit

8 INUS steht für *Insufficient but necessary part of an unnecessary but sufficient condition*.

dieser Einflüsse wird als ‚Mechanismen‘ bezeichnet, sie erklären das Eintreten eines geplanten Wirkungspfads – oder Abweichungen davon.

Mechanismen wirken nicht immer und überall gleich, sondern in Abhängigkeit vom jeweiligen Kontext. Dieser umfasst die institutionellen, organisatorischen oder sozialen Rahmenbedingungen, die die konkrete Ausprägung der Mechanismen beeinflussen. Sie werden wiederum zu Gruppen von Bedingungen gebündelt, die für einen bestimmten Kontext charakteristisch sind. Die Auswahl hängt eng mit der zu untersuchenden Situation zusammen und kann auf unterschiedlichen Ebenen angesiedelt sein: Bei der Herstellung einer Speise geht es eher um persönliche (z. B. Erfahrung, Interesse) und infrastrukturelle Faktoren (z. B. verfügbarer Herd und Küchengeräte). Bei einem Ausbildungsvorhaben sind auch Rahmenbedingungen auf der Makroebene relevant, wie z. B. kulturelle Faktoren (z. B. der Stellenwert von Bildung) oder nationale Politiken (z. B. die Standorte von Schulen oder die Ausbildung von Lehrkräften).

BEISPIELE FÜR GEEIGNETE ANSÄTZE ZUR KAUSALANALYSE

Nachfolgend werden drei Ansätze skizziert, die bei komplexen und komplizierten Zusammenhängen angewendet und den obigen Kausalitätsmustern zugeordnet werden können. Sie sind für die Analyse von Wirkungen auf Systeme geeignet.

LOGISCHE STRATEGIEN

Beim Zusammenwirken mehrerer Faktoren können verschiedene **logische Strategien** angewendet werden, um zu überprüfen, ob die Faktoren jeweils notwendig und hinreichend waren. Die wichtigsten Strategien für die Untersuchung des Zusammenhangs von Intervention und Wirkung sind (vgl. DAVIDSON 2005):

- » **Zeitliche Abfolge:** Die Wirkung (Y) tritt immer erst nach der dafür maßgeblichen Intervention (X) auf, z. B. nach den Leistungen/Aktivitäten eines Vorhabens. Dabei sollte aber auch der Zeitbedarf für das Auftreten eines Effekts berücksichtigt werden.
- » **Übereinstimmung mit Dosis:** Die Wirkung (Y) korreliert mit dem Ausmaß der Intervention (X), diese Beziehung muss jedoch nicht linear sein (es können z. B. Schwellen- oder Grenzwerte auftreten). Eine Variante ist, wenn Y in unterschiedlichen Kontexten immer in Verbindung mit X auftritt.

- » **Inhaltliche Übereinstimmung:** Die beobachtete Wirkung (Y) entspricht inhaltlich der dafür maßgeblichen Intervention (X). Um Aufschluss über andere Einflussfaktoren zu bekommen, sollte gezielt nach gegenteiligen Beispielen gesucht werden, in denen z. B. dieselbe Wirkung durch eine alternative Intervention erzielt wurde.
- » **Assoziierung:** Die Wirkung Y tritt immer (bzw. häufig) in Zusammenhang mit der Intervention (X) auf, oder in Verbindung mit einem dritten Faktor (Z).
- » **Ausschlussverfahren:** Die Wirkung (Y) wird als Folge einer Intervention (X) angesehen, da andere mögliche Faktoren nicht beobachtet werden.

CONTRIBUTION ANALYSIS

Diese wurde für Kausalitäten des INUS-Typs entwickelt (vgl. MAYNE 2011). Auf eine direkte Zuschreibung von Wirkungen wird verzichtet, stattdessen wird der Beitrag (*contribution*) eines Vorhabens zu den intendierten Effekten und den Fortschritten bei ihrer Erzielung erfasst. Dadurch kann begründet werden, dass ein Vorhaben notwendiger Bestandteil einer Strategie ist, die ausreicht um – in Kombination mit anderen Faktoren – die angestrebte Veränderung zu erreichen; oder welche Vorteile eine Strategie unter unterschiedlichen Rahmenbedingungen bzw. gegenüber anderen möglichen Strategien hat (zur bisherigen Anwendungspraxis siehe insbesondere Journal Evaluation 18, Issue 3).

Auf der Basis einer Handlungstheorie des Vorhabens und von Daten über die erzielten Effekte wird die Plausibilität der Beiträge wie folgt herausgearbeitet:

- » Aufbereitung der Beiträge (*contribution story*), v. a. durch eine Analyse der Informationen entlang der Wirkungskette. Die Analyse wird anhand der Meinungen unterschiedlicher *Stakeholders* überprüft; bei hoher Übereinstimmung der Meinungen ist die Analyse tendenziell verlässlicher.
- » Überprüfung der Beiträge durch den Ausschluss alternativer Erklärungen für das Auftreten der einzelnen Effekte.
- » Falls diese Überprüfung Lücken aufdeckt: Suche nach zusätzlichen Informationen für die betreffenden Beiträge und darauffolgende Überarbeitung der *contribution story*.

REALIST EVALUATION

Dieser Ansatz basiert auf dem generativen Kausalitätsverständnis (vgl. PAWSON, TILLEY 1997). Sein Ziel ist es, das Eintreten (oder Unterbleiben) von Wirkungen durch die Mechanismen zu erklären, mit denen sie ausgelöst werden (sollen). Die Leitfrage lautet: „Was wirkt für wen unter welchen Rahmenbedingungen?“ Damit können besonders fundierte Erklärungen für Wirkungsweisen geliefert werden. Allerdings erfordert dies detaillierte Kenntnisse der Situation, die Berücksichtigung vielfältiger Faktoren und eine vollständige Beschreibung der Wirkungspfade (was wiederum ausreichend Zeit und Ressourcen voraussetzt).

Die wichtigsten Eckpunkte bei der Analyse sind:

- » Die Untersuchung der Zusammenhänge zwischen den Leistungsprozessen eines Vorhabens (Mechanismen), den Rahmenbedingungen (Kontext) und den beobachteten Effekten (outcomes).
- » Die Ermittlung der Wirkungsweise über die Identifizierung und Erklärung von Mustern („Kontext – Mechanismus – Outcome Konfigurationen“). Diese Muster können in Form einer Realist Matrix übersichtlich dargestellt werden.
- » Die Aufstellung (vorab) von Hypothesen über die erwarteten Konfigurationen. Das (Nicht)Zutreffen der Hypothesen wird während oder nach der Implementierung analysiert.

Nach dieser Betrachtung systemischer Herangehensweisen zur Analyse von Wirkungen und Kausalität wendet sich das abschließende Kapitel wieder praktischen Überlegungen zu: den erforderlichen Rahmenbedingungen und dem Mehrwert der Anwendung systemischer Ansätze.

5 MEHRWERT UND RAHMENBEDINGUNGEN SYSTEMISCHER ANSÄTZE

In der internationalen Fachdiskussion besteht weitgehend Übereinstimmung, dass kein Evaluationsansatz den Anspruch eines *Gold Standards* erheben sollte. Ein Evaluationsansatz sollte v. a. der jeweiligen Situation und den Anforderungen gerecht werden. Auch eine systemische Herangehensweise ist nicht immer angebracht und sollte nur dann gewählt werden, wenn von ihr ein besonderer Nutzen zu erwarten ist und geeignete Rahmenbedingungen bestehen.

5.1 Potentiale und Grenzen

Die Vorteile und Potentiale systemischer Ansätze in der Evaluation können in vier Punkten zusammengefasst werden:

ERWEITERUNG DES HANDLUNGSREPERTOIRES BEI KOMPLEXEN EVALUATIONSGEGENSTÄNDEN

Der Einsatz systemischer Ansätze ist in komplizierten Situationen **nützlich**, in die z. B. unterschiedliche Akteure (mit teilweise divergierenden Ansichten) bzw. Ebenen involviert sind oder in denen mehrere Wirkungspfade auftreten (parallel oder alternativ). Systemische Methoden stehen z. B. für die Visualisierung dynamischer Beziehungen, die Berücksichtigung unterschiedlicher Perspektiven, oder die Unterscheidung zwischen Vorhaben und Kontext zur Verfügung.

Der Einsatz systemischer Ansätze ist **unabdingbar** in komplexen Situationen, in denen Wechselwirkungen oder emergente Phänomene auftreten. Dies ist in nahezu allen Projekten der EZ der Fall. Dadurch können passende Methoden für die Modellierung nicht-linearer Beziehungen oder die Erfassung von Mustern vielfältig verbundener Elemente eingesetzt werden.

ANGEMESSENE VEREINFACHUNG VON SITUATIONEN ZUNEHMENDER KOMPLEXITÄT

Mit Systemischen Ansätzen wird linear-kausales Denken durch eine zirkuläre Betrachtungsweise ersetzt; dadurch werden bei der Reduktion von Komplexität die Eigenschaften und Prinzipien sozialer Systeme berücksichtigt. Systemische Ansätze erlauben es, Situationen sowohl in ihrer Gesamtheit zu betrachten als auch bei Bedarf in Subsysteme (als Teile eines größeren Ganzen) zu zerlegen. Ferner werden die Wechselwirkungen der betreffenden Systeme (z. B. zwischen Vorhaben und Zielsystem) berücksichtigt. Dies erhöht die Kontextorientierung von Evaluationen, denn die Betrachtungseinheit ist immer das Vorhaben im jeweiligen Kontext.

ÜBERPRÜFUNG DER ANSPRÜCHE VON SYSTEMVERÄNDERUNGEN

Vorhaben der Entwicklungszusammenarbeit haben häufig explizite Ansprüche an die Veränderung größerer Systeme. Mit systemischen Ansätzen lassen sich die erhofften Impulswirkungen über mehrere Ebenen und Bereiche passend modellieren und beobachten. Wo die Veränderung bestehender Systeme Teil der beabsichtigten Entwicklung ist, können systemische Ansätze zum Verständnis dieser (sozialen) Systeme und zur Identifizierung von Hebeln beitragen.

„SENSE-MAKING“ IN SCHWIERIGEN SITUATIONEN

Systemische Ansätze liefern Antworten darauf, warum eine Intervention im konkreten Kontext funktioniert (oder auch nicht). Ihr Einsatz bietet sich an, wenn in Evaluationen Sachverhalte auftreten, die mit herkömmlichen Methoden nicht – oder nicht ausreichend – analysiert und erklärt werden können. Dies ist beispielsweise der Fall, wenn die Daten Widersprüche oder Überraschungen aufweisen, unerklärliche Phänomene auftreten oder Akteure ein unerwartetes Verhalten zeigen. In derartigen Fällen können systemische Ansätze anlassbezogen eingesetzt werden und müssen nicht von vorneherein geplant sein.

Die Anwendung systemischer Ansätze hat aber auch **Nachteile bzw. Grenzen**. Systemkonzepte und daraus abgeleitete Methoden können relativ abstrakt wirken, was ihre Vermittlung unter Umständen erschwert. Einige Methoden sind in Bezug auf Zeit und Kosten aufwendig – jedoch können viele systemische Methoden ohne großen Aufwand eingesetzt werden und funktionieren relativ rasch, z. B. bei der Darstellung von Beziehungen und der Berücksichtigung verschiedener Perspektiven.

Was die Haltung der Beteiligten betrifft, so erfordert systemisches Arbeiten mitunter Flexibilität seitens des Auftraggebers. Wichtig ist insbesondere ein Verständnis für Evaluation als iterativen Prozess, wobei ihr Design zunächst nur grob festgelegt wird und für neue Anforderungen und Fragestellungen offen bleibt.

5.2 Rahmenbedingungen für systemische Ansätze

Die gewinnbringende Anwendung systemischer Ansätze hängt maßgeblich von den jeweiligen Rahmenbedingungen von Evaluationen ab.

FÖRDERLICHE RAHMENBEDINGUNGEN

Systemische Ansätze eignen sich vorzüglich für anwendungsorientierte Evaluationen, deren NutzerInnen bekannt und an allen Phasen der Evaluation beteiligt sind. Dies gilt auch für Evaluationen, bei denen Reflexion und gemeinsames Lernen im Vordergrund stehen und Kontroll- bzw. Rechenschaftsaspekte entweder keine Rolle spielen oder so eingebettet sind, dass sie mit diesem Lernfokus kompatibel bleiben.

AuftraggeberInnen und AkteurInnen im evaluierten System sollten für die Infragestellung von Werten und Handlungsmustern Spielraum geben. Systemische Ansätze entfalten ihre Wirksamkeit besonders da, wo die Agierenden im evaluierten System offen und bereit sind für einen konstruktiven Umgang mit Unterschieden oder Konflikten.

Seitens der EvaluatorInnen ist selbstverständlich die Kompetenz zur Anwendung systemischer Methoden erforderlich – sowie die Fähigkeit, auf flexible und konstruktive Weise unterschiedliche Rollen (z. B. auch ModeratorIn, Coach, Impulsgebende) einzunehmen, sofern dies zweckdienlich ist und mit den Beteiligten vereinbart wurde.

HINDERLICHE RAHMENBEDINGUNGEN

Einige Haltungen und Situationen können den Einsatz systemischer Ansätze behindern: EvaluatorInnen, die sich vorrangig als unabhängig ‚Prüfende‘ sehen und nicht bereit oder imstande sind, andere Aufgaben oder Haltungen anzunehmen, dürften bei der Anwendung systemischer Ansätze Mühe haben. Seitens der Auftraggeber und der Akteure im evaluierten System kann eine geringe Kritikfähigkeit den Einsatz systemischer Methoden hemmen.

Wenn eine Evaluation primär als externe Kontrolle wahrgenommen wird (z. B. wenn sie mit Entscheidungen für oder gegen eine Vertragsverlängerung und mit existenziellen Konsequenzen für die Beteiligten einhergeht), kann es schwierig sein, die für systemische Ansätze notwendige Offenheit zu schaffen. Ebenso setzen AuftraggeberInnen, die einen klaren Bedarf nach externer, ‚objektiver‘ Prüfung bzw. Legitimation durch unabhängige ExpertInnen signalisieren, den Entfaltungsmöglichkeiten systemischer Ansätze enge Grenzen. Dies gilt auch für Situationen, in denen wenig Handlungsspielraum bei der Implementierung der Evaluation besteht – z. B. durch knapp bemessene Zeitpläne, oder detaillierte methodische Vorgaben in Ausschreibungen und Auftragsrahmen (TOR), die systemische Ansätze ausschließen.

LITERATURVERZEICHNIS

- BAGADION, B. C. JR. (1999): Bending the Wind. Lessons from Mt. Makiling. Empowering People for Natural Resource Management. AIM Center for Development Management: Makati City.
- BAUMFELD, L.; HUMMELBRUNNER, R.; LUKESCH, R. (2009): Instrumente systemischen Handelns. Leonberg: Rosenberger Fachverlag.
- CAPPER, P. (2004): Contradiction Analysis. (abrufbar unter URL: <http://www.webresearch.co.nz>, Stand: März 2010).
- DAVIDSON, J. (2005): Evaluation Methodology Basics. The Nuts and Bolts of Sound Evaluation. London: Sage.
- DFID (2012): Broadening the range of Designs and Methods for Impact Evaluations, Working Paper 38. London.
- DICK, B. (2002): Convergent Interviewing. (abrufbar unter URL: http://www.uqconnect.net/action_research/arhome.html, Stand: März 2010).
- EARL S.; CARDEN F.; SMUTYLO T. (2001): Outcome Mapping: Building Learning and Reflection into Development Programs. Ottawa: IDRC.
- FOERSTER, H. VON (1999): Sicht und Einsicht. Heidelberg: Carl-Auer-Systeme.
- FUNNELL, S.; PATRICIA, C.; ROGERS, J. (2011): Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. San Francisco, CA: John Wiley & Sons.
- GLOUBERMAN, S.; ZIMMERMAN, B. (2002): Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like? Commission on the Future of Health Care in Canada, Discussion Paper 8.
- GTZ (2008): Capacity Works. Das Managementmodell für nachhaltige Entwicklung. Eschborn.

- HABERFELLNER, R. ET AL. (1999): Systems Engineering – Methodik und Praxis. Zürich: Verlag Industrielle Organisation.
- HORN, S. (2011): Wirkungsorientierte Evaluation nichtstaatlicher deutscher bilateraler Berufsbildungszusammenarbeit in Ghana. Entwicklung eines systemischen Evaluationsverfahrens im Rahmen einer Fallstudie des vom Evangelischen Entwicklungsdienst e. V. unterstützten Programms Vocational Training for Females. Dresden. (abrufbar unter URL: <http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-77576>, Stand: Juli 2013).
- JACKSON, P. Z.; MCKERGOW, M. (2002): Solution Focus. The SIMPLE Way to Positive Change. London: Nicolas Brealey.
- JANSEN, D. (2006): Einführung in die Soziale Netzwerkanalyse. Grundlagen, Methoden, Forschungsbeispiele. 3. Auflage. Wiesbaden: Springer.
- JOURNAL EVALUATION (2012): Special Issue: Contribution Analysis. July 2012:18 (3). SAGE.
- KURTZ, C. F.; SNOWDEN, D. J. (2003): The new dynamics of strategy: Sense-making in a complex and complicated world. In: IBM Journal of Research and Development 42, No. 3, S. 462–483.
- MALIK, F. (1984): Strategie des Managements komplexer Systeme. Ein Beitrag zur Management-Kybernetik evolutionärer Systeme. Bern: Haupt.
- MAYNE, J. (2011): Contribution Analysis: Addressing Cause and Effect. In: Forss, K., Marra, M.; Schwartz, R. (2011): Evaluating the Complex: Attribution, Contribution, and Beyond. New Jersey: Transaction Press.
- O'CONNOR, J.; MCDERMOTT, I. (2000): Die Lösung lauert überall. Systemisches Denken verstehen und nutzen. Kirchzarten bei Freiburg: VAK Verlag.
- PATTON, M. Q. (2010): Developmental evaluation. New York: Guilford Press.
- PROBST, G. J.B.; ULRICH, H. (1988): Anleitung zum ganzheitlichen Denken und Handeln. Bern: Haupt
- PAWSON, R.; TILLEY, N. (1997): Realistic Evaluation. Thousand Oaks: Sage.
- PAWSON, RAY (2006): Evidence-based Policy: A Realist Perspective. London: Sage.

- REYNOLDS, M.; HOLWELL, S. EDS. (2010): *Systems Approaches to Managing Change a practical guide*. London: Springer Press.
- ROGERS, P. (2008): Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. In: *Evaluation* 14 (1): S. 29–48.
- STOCKMANN, R. (2006): *Evaluation und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement*. Waxmann: Münster.
- ULRICH, W.; REYNOLDS, M. (2010): Critical Systems Heuristics. In: Reynolds, M.; Holwell, S. eds. (2010): *Systems Approaches to Managing Change: a practical guide*. London: Springer Press.
- VESTER, F. (1999): *Die Kunst, vernetzt zu denken – Ideen und Werkzeuge für einen neuen Umgang mit Komplexität*. Stuttgart: Deutsche Verlags-Anstalt.
- WILLIAMS, B.; HUMMELBRUNNER, R. (2011): *Systems Concepts in Action. A Practitioner's Toolkit*. Stanford: Stanford University Press.
- WILLIAMS, B.; IMAM, I. EDS. (2007): *Systems concepts in evaluation: An expert anthology*. Point Reyes: American Evaluation Association/EdgePress.

**„OBJEKTIV GEHT NICHT MEHR.“ –
EXPERTENINNENSICHT AUF SYSTEMISCHE
EVALUATIONSANSÄTZE IN DER
ENTWICKLUNGSZUSAMMENARBEIT**

Anja Bugenhagen

1 EINLEITUNG

In der internationalen wissenschaftlichen Debatte über Evaluation wird deutlich, dass bestimmte Evaluationsansätze zeitweise einen gewissen ‚Hype‘ erfahren und aus unterschiedlichen Gründen besonders im Fokus des Interesses stehen. Wie unter anderem in den Artikeln dieses Bandes aufgezeigt, werden aktuell systemische Evaluationsansätze diskutiert und in der Evaluationsforschung relevant gemacht (vgl. auch WILLIAMS, IMAM 2006; PATTON 2010; HUMMELBRUNNER 2011), wobei sich diese Debatte, zumindest in Deutschland, noch in ihren Anfängen befindet. Dieser Beitrag möchte etwas zu der Fachdiskussion beisteuern und rückt in den Mittelpunkt, inwiefern EvaluatorInnen und WissenschaftlerInnen systemische Evaluationsansätze charakterisieren bzw. von anderen abgrenzen und darüber hinaus die Chancen und Grenzen dieser ‚neuen‘ Herangehensweisen einschätzen. Das Ziel dieses Beitrages besteht vorrangig darin, den Kommunikationsprozess über systemische Evaluationsansätze zu reflektieren und weniger, eine wissenschaftlich fundierte Charakterisierung oder Nutzen-Kosten-Analyse dieser scheinbar neuen Evaluationsansätze zu liefern.

Um das wachsende Interesse an systemischen Evaluationsansätzen zu verstehen, ist es zunächst sinnvoll, sich die Herausforderungen zu vergegenwärtigen, denen sich Evaluation speziell in der Entwicklungszusammenarbeit stellen muss. Unter anderem werden in diesem Beitrag die in der jüngsten Vergangenheit vollzogenen Veränderungen in der inhaltlichen Ausrichtung und Organisation der deutschen und internationalen Entwicklungszusammenarbeit kurz beleuchtet, um nachvollziehbar zu machen, warum fortwährend die Frage nach adäquaten Evaluationsmethoden und -modellen gestellt wird. Während in einem Zwischenfazit die (neuen) Anforderungen und Erwartungen an Evaluation kurz zusammengefasst werden, wird in den darauffolgenden Abschnitten dargelegt, warum insbesondere systemische Ansätze in der Evaluation an Bedeutung gewinnen können bzw. aktuell stark diskutiert werden. Auf der Grundlage von fünf ExpertInnen-Interviews werden Verständnisse und mögliche Charakterisierungen von systemischen Evaluationsansätzen herausgestellt. Ein detaillierter Blick auf die Aussagen und Erfahrungen einzelner InterviewpartnerInnen ermöglicht darüber hinaus eine Markierung möglicher Chancen und Grenzen systemischer Evaluationsansätze. Abschließend werden die wichtigsten Erkenntnisse zusammengetragen.

2 HERAUSFORDERUNGEN VON EVALUATIONEN IN DER ENTWICKLUNGSZUSAMMENARBEIT

2.1 Rahmenbedingungen der Entwicklungszusammenarbeit – Wandel, Millennium, Legitimation

Zu Beginn des 21. Jahrhunderts steht die internationale Entwicklungszusammenarbeit unter einer noch nie dagewesenen wissenschaftlichen und medialen Kritik. Bezüglich der intendierten globalen Armutsreduktion zeigten sich nur bescheidene Fortschritte (ausgenommen China, Indien und die Tigerstaaten), gleichwohl lokal und punktuell durchaus positive Ergebnisse erzielt wurden (vgl. BORRMANN, STOCKMANN 2009a: 34). Auch wenn die Kritik an der Entwicklungszusammenarbeit so alt ist wie die Entwicklungszusammenarbeit selbst (vgl. STOCKMANN et al. 2010: 459), mehrten sich um die Jahrtausendwende Stimmen, welche nicht nur die Qualität und Wirksamkeit in Frage stellten, sondern öffentlich die Abschaffung der Entwicklungszusammenarbeit forderten (vgl. KEVENHÖRSTER, VAN DEN BOOM 2009: 79 f.; NUSCHELER 2004: 13 ff.). Da diese fundamentale Kritik bis heute nicht abebbt und in zyklischen Abständen öffentliche Diskurse prägt, steht die internationale und damit ebenso die deutsche staatliche Entwicklungszusammenarbeit unter einem enormen Legitimationsdruck. Dieser Druck wird durch Finanz- und Wirtschaftskrisen noch zusätzlich erhöht und bewirkt eine stärkere Bindung des Mittelflusses an die Erfüllung bestimmter Konditionalitäten. Im Lichte knapp bemessener Gelder für die Entwicklungszusammenarbeit, ist daher mehr denn je ein möglichst effizientes Agieren gefragt und gefordert (vgl. BORRMANN, STOCKMANN 2009a: 34). Dies hat letztlich dazu geführt, dass die Entwicklungszusammenarbeit das einzige Politikfeld in Deutschland darstellt, welches über ein ausgefeiltes Monitoring- und Evaluationssystem verfügt (vgl. STOCKMANN et al. 2010: 459).

Dass das Erzielen sichtbarer und vor allem nachweisbarer Erfolge seit der Jahrtausendwende im Mittelpunkt steht, kommt nicht zuletzt auch in den von den Vereinten Nationen (UN) verabschiedeten ‚Millenniumsentwicklungszielen‘ (englisch: Millennium Development Goals, MDGs) zum Ausdruck (vgl. Tab. 2). Sie konstituie-

ren seit 2000 den Rahmen der internationalen Entwicklungszusammenarbeit und stellen konkrete Ziele für die globale Armutsreduktion dar, welche bis zum Jahr 2015 verwirklicht werden sollen. Die acht MDGs werden dabei durch Zielvorgaben konkretisiert und durch „Indikatoren zur Überwachung der Fortschritte“ (UN 2005: 1) messbar gemacht. Die Millennium-Deklaration und die daraus resultierenden MDGs stehen somit repräsentativ für eine ganze Reform-Dynamik, die im Sinne einer Effizienz- und Effektivitätssteigerung der internationalen Entwicklungszusammenarbeit in Gang gesetzt wurde (vgl. BORRMANN, STOCKMANN 2009a: 35 f.).

Ziel 1	den Anteil der Weltbevölkerung, der unter extremer Armut und Hunger leidet, halbieren
Ziel 2	allen Kindern eine Grundschulbildung ermöglichen
Ziel 3	die Gleichstellung der Geschlechter fördern und die Rechte von Frauen stärken
Ziel 4	die Kindersterblichkeit verringern
Ziel 5	die Gesundheit der Mütter verbessern
Ziel 6	HIV/AIDS, Malaria und andere übertragbare Krankheiten bekämpfen
Ziel 7	den Schutz der Umwelt verbessern
Ziel 8	eine weltweite Entwicklungspartnerschaft aufbauen

Tab. 1: Die acht Milleniumsentwicklungsziele; *Quelle: BMZ 2010: Vorwort*

Eine weitere Initiative, welche die Entwicklungszusammenarbeit insgesamt und die Evaluation der Entwicklungszusammenarbeit im Speziellen nachhaltig beeinflusst, bildet die Paris-Deklaration. Während 2003 in Rom ein erstes hochrangiges Forum zur Wirksamkeit der Entwicklungszusammenarbeit organisiert wurde, fand im Frühjahr 2005 eine Folgekonferenz in Paris statt, aus der die Paris-Deklaration mit fünf Grundprinzipien zur wirksamen Zusammenarbeit hervorging (vgl. BMZ 2005):

- » Die Eigenverantwortung der Entwicklungsländer soll gestärkt werden (,Ownership‘).
- » Die Geber sollen Institutionen der Kooperationsländer nutzen und ihre Programme an den Strategien und Verfahren der Partner ausrichten (,Alignment‘).
- » Die Geber sollen ihre Programme und Verfahren untereinander abstimmen und harmonisieren (,Harmonisation‘).
- » Die Maßnahmen sollen auf Ergebnisse ausgerichtet sein (,Managing for Results‘). Das heißt, die Geber lassen sich an den Ergebnissen

ihres entwicklungspolitischen Handelns messen (z. B. Verringerung der Analphabetenquote) und nicht am finanziellen Einsatz (z. B. zehn Millionen Euro für neue Schulen).

- » Geber- und Kooperationsländer sollen gemeinsam über ihr entwicklungspolitisches Handeln gegenüber der Öffentlichkeit und den Parlamenten Rechenschaft ablegen („Accountability“) (BMZ 2005).

Ebenso an dieser Stelle wird durch die verbindliche Festlegung von zwölf Fortschrittsindikatoren mit dazugehörigen Zielvorgaben, welche bis 2010 erreicht werden sollten, die Entwicklungszusammenarbeit einem beträchtlichen Umsetzungsdruck ausgesetzt. Dabei ist der Paris-Deklaration die Hypothese zugrunde gelegt, dass durch eine bessere Koordination und Harmonisierung die Wirkung der Entwicklungszusammenarbeit gesteigert werden kann und gleichzeitig Transaktionskosten dezimiert werden können. Die Fusion der drei großen BMZ-Durchführungsorganisationen GTZ, DED und InWEnt zur GIZ steht daher ganz im Zeichen der Paris-Deklaration. Generell bilden diese fünf Grundprinzipien von da an einen wichtigen Bezugsrahmen für die Evaluation der Entwicklungszusammenarbeit, indem sie vor allem Evaluation als verpflichtendes Element zur Rechenschaftslegung einen großen Stellenwert einräumen und eine konsequente Ausrichtung an Entwicklungsergebnissen vorgeben (vgl. BORRMANN, STOCKMANN 2009a: 35 f.).

2.2 Veränderte Umsetzungsmodalitäten

Korrespondierend mit den angesprochenen internationalen Vereinbarungen und Erwartungen, haben sich sowohl das Selbstverständnis als auch die Umsetzungsmodalitäten der Entwicklungszusammenarbeit stark gewandelt. Demnach versteht sich die Entwicklungszusammenarbeit heute primär als Makro-Strukturpolitik, welche hinsichtlich einer intendierten Armutsminderung auf eine Verbesserung ökonomischer, ökologischer, politischer und institutioneller Rahmenbedingungen auf globaler und nationaler Ebene abzielt. Die Lehre aus über 60 Jahren Entwicklungspolitik hat die Perspektive evoziert, dass die Kontextbedingungen und das Entwicklungsniveau der einzelnen Länder ausschlaggebend dafür sein müssen, wie Entwicklung vorangetrieben wird und welche Konzepte fruchten. Patentrezepte, die allen übergestülpt werden können, gibt es nicht. Auf diese Erfahrungen aufbauend betrachtet die Entwicklungszusammenarbeit ihre Entwicklungsprogramme als von außen induzierte Interventionen in ein aus Sicht der Geberländer unbekanntes und unverständliches Empfängerland. Daher wird eine an die jeweilige Situation angepasste Mischung von Entwicklungsprinzi-

prien und Instrumenten als erfolgversprechend bewertet, um eine Entwicklung ‚von innen‘ zu bewirken (vgl. STOCKMANN et al. 2010: 374; RAUCH 2012: 50 f.; GTZ 2010: 22 ff.).

Dieser beschriebene Paradigmenwechsel in der Entwicklungszusammenarbeit spiegelt sich unmittelbar in der konzeptionellen Ausrichtung sowie in den Umsetzungsmodalitäten wider. So gewinnen generelle und sektorale Budgethilfen an Bedeutung, während hingegen isolierte Projekte an Gewicht verlieren, sodass heutzutage unter dem Begriff der Entwicklungszusammenarbeit eine Vielzahl an mehr oder weniger komplexen Mischformen von Projekt-, Programm- und Budgethilfen subsumiert wird. Die bunte Mischung der Durchführungsformen z. B. in Gestalt einer parallelen Umsetzung von bi- und multilateralen Ko-Finanzierungen, Programmhilfen, Korbfinanzierungen mit den wichtigsten AkteurInnen eines Sektors und generellen Budgethilfen ist mittlerweile die Regel (vgl. BORRMANN, STOCKMANN 2009a: 37). Nachdem sich die Entwicklungszusammenarbeit vorrangig eindimensionaler, auf einzelne Schlüsselfaktoren, Konzepte und Instrumente stützende Ansätze bediente und dieses Vorgehen letztlich nicht die erhofften Erfolge brachte, rücken angesichts der Komplexität entwicklungspolitischer Aufgaben und Ziele nun zunehmend multidimensionale Strategien in den Vordergrund (vgl. RAUCH 2012: 46 ff.; GTZ 2010: 14). Die jahrelangen Erfahrungen offenbarten den Erkenntnisraum, dass Armut als ein Ergebnis multifaktorieller Entwicklungsprozesse zu verstehen ist und damit eindimensionale Betrachtungsweisen den realen Zusammenhängen nicht gerecht werden. Multidimensionale Strategien tragen dieser Perspektive Rechnung, indem eine sektorenübergreifende Koordination der Implementierung von Entwicklungsprogrammen angestrebt wird, was das Zusammenarbeiten multidisziplinärer Teams von Fachleuten erforderlich macht. Diese neue Betrachtungs- und Vorgehensweise stößt jedoch an Grenzen praktischer Umsetzbarkeit angesichts sektoral organisierter Verwaltungen und ist nur im Rahmen spezieller Sonderorganisationen machbar (vgl. RAUCH 2012: 49). Gleichwohl wird dem übersektoralen Denken und Handeln in der Entwicklungszusammenarbeit eine wachsende Bedeutung zuteil wie das folgende Zitat verdeutlicht:

„Während zu Beginn und Mitte der 1990er Jahre punktuelle Maßnahmen mit einem Krisenbezug entwickelt wurden und dieser Ansatz in der zweiten Hälfte der 1990er Jahre zu einem Konzept zur zivilen Krisenprävention weiterentwickelt wurde, stehen mittlerweile übergreifende Ansätze im Mittelpunkt, mit deren Hilfe Beiträge zu Stabilität und Sicherheit geleistet werden sollen. Dabei geht es vor allem um ein stärker vernetztes Denken und Handeln zwischen den sicherheits-, außen- und entwicklungspolitischen Akteuren“ (KLINGEBIEL 2013: 36).

Als Folge dieser strategischen Neuausrichtung nimmt die Anzahl der an einem Vorhaben beteiligten AkteurInnen stark zu und setzt andererseits ein noch höheres Maß an Koordination in der ohnehin von einer Vielfalt von staatlichen und privaten AkteurInnen gekennzeichneten institutionellen Landschaft voraus (vgl. RAUCH 2012: 46 ff.). All dies mündet letztlich in einem drastischen Anstieg der Komplexität von Maßnahmen in der Entwicklungszusammenarbeit (vgl. DEVAL 2012).

In Anbetracht der Tatsache, dass sich die Ziele und die Instrumente der Entwicklungszusammenarbeit sowohl auf der nationalen als auch auf der internationalen Maßstabsebene stark gewandelt haben und zudem die Anforderungen zur Rechenschaftslegung enorm gestiegen sind, werden insbesondere Evaluationssysteme vor ganz neue Herausforderungen gestellt (vgl. BORRMANN, STOCKMANN 2009a: 31). Eben diese Herausforderungen sollen in den folgenden Kapiteln vertieft werden.

2.3 Wirkungsmessung – oder wie lässt sich erfolgreiche Entwicklungszusammenarbeit darstellen?

KOMPLEXITÄT UND GRENZEN DER WIRKUNGSMESSUNG

Wie bereits in vorangegangenen Abschnitten beleuchtet, befindet sich das gesamte System der Entwicklungszusammenarbeit durch die entwicklungspolitischen Agenden im Umbruch. Anhand der Millenniumsentwicklungsziele und der Paris-Deklaration kristallisiert sich deutlich heraus, dass verstärkte Anstrengungen zum glaubwürdigen Wirkungsnachweis der Entwicklungszusammenarbeit offiziell gefordert werden. Um diesen Erwartungen nachzukommen, wird eine noch konsequentere Steuerung nach Resultaten fokussiert über die Konstituierung wirkungsorientierter Monitoring- und Evaluationssysteme. Dennoch wird die Wirkungsmessung trotz großer Anstrengungen und hoher Ambitionen bislang als defizitär eingeschätzt (vgl. BORRMANN, STOCKMANN 2009a: 46). Die Gründe hierfür sind mannigfaltig:

Es existiert eine Reihe an methodischen Herausforderungen, der sich die Wirkungsanalyse stellen muss. Wird unter dem Begriff ‚Wirkung‘ ein veränderter oder erhaltener Zustand verstanden, dem Ursachen zur Entstehung zugrundeliegen und welcher eben ohne jene Ursachen nicht zustande gekommen wäre, so kann jede Maßnahme in der Entwicklungszusammenarbeit unterschiedlichste Wirkungen erzeugen. Diese können in ihrer Gesamtheit und als sich gegenseitig beeinflussendes Wirkungsgeflecht in der Praxis weder vollständig erfasst und beschrieben noch gemessen werden. In diesem

Zusammenhang sei neben zahlreichen methodischen Hürden in der Wirkungsanalyse als eine zentrale das Zuordnungsproblem erwähnt:

„Das zentrale Problem bei der Wirkungsanalyse ist der Nachweis, dass bestimmte Veränderungen – oder auch ‚konservierende Wirkungen‘, d. h. Wirkungen, die im Erhalt eines Ausgangszustands bestehen – kausal mit der zu untersuchenden Entwicklungsmaßnahme zusammenhängen. Selbst wenn festgestellt werden kann, dass die angestrebten Veränderungen eingetreten sind, können andere Faktoren dafür verantwortlich sein“ (ARBEITSKREIS EVALUATION VON ENTWICKLUNGSPOLITIK DER DEGEVAL 2009: 14).

Auf der Grundlage der Zuordnungsproblematik wird bei der Wirkungsmessung häufig auf eine größere Bandbreite an verschiedenen Nachweis- und Vergleichsmethoden zurückgegriffen. In der quantitativen Sozialforschung werden insbesondere experimentelle Untersuchungen mit Kontrollgruppen als Königsweg der Wirkungsanalyse gepriesen. Diese ermöglichen es, mit Hilfe von soliden Kosten-Wirkungsanalysen, verschiedene Maßnahmen auf ihre Effizienz und Effektivität hin zu überprüfen. Jedoch impliziert auch dieses methodische Vorgehen große praktische und akademische Herausforderungen, um zu einer Bildung vergleichbarer Kontrollgruppen zu gelangen. So müsste bereits zu Beginn der Entwicklungsmaßnahme eine genaue Planung der Wirkungsanalyse mitgedacht werden, was den mittlerweile sehr breit gefächerten Inhalten der Entwicklungsmaßnahmen und somit der Vielfalt der anvisierten Wirkungen entgegensteht und überdies auch andere entwicklungspolitische Grundsätze konterkarieren kann (z. B. der aktiven Partizipation der Zielgruppen bei der Durchführung der Entwicklungsmaßnahme). Insgesamt erweist sich diese Herangehensweise somit in ihrer Anwendung in der Entwicklungszusammenarbeit in vielen Fällen aus praktischen und ethischen Gründen als nicht umsetzbar oder als äußerst aufwändig (vgl. ARBEITSKREIS EVALUATION VON ENTWICKLUNGSPOLITIK DER DEGEVAL 2009: 13 ff.):

„Sollen Kontrollgruppen bei einer Wirkungsanalyse (z. B. im Rahmen einer Evaluierung) einbezogen werden, müssen solche zunächst gefunden oder konstruiert werden. Sie müssen unter den gleichen Rahmenbedingungen leben und von den gleichen Voraussetzungen ausgehen wie die Zielgruppen, was oft zusätzliche Untersuchungsschritte erfordert. Je größer und komplexer Programme sind im Hinblick auf alle Kriterien und die anvisierten Wirkungen, umso schwieriger ist dies zu realisieren und manchmal ist es unmöglich“ (ARBEITSKREIS EVALUATION VON ENTWICKLUNGSPOLITIK DER DEGEVAL 2009: 18).

Wie das Zitat unter anderem anklingen lässt, spielen ebenso die entwicklungs- politischen Spezifika der Maßnahmen und Programme bei der Wirkungsanalyse eine nicht zu unterschätzende Rolle. Demzufolge wird die ohnehin vorhandene Komplexität sowie der Schwierigkeitsgrad der Wirkungsmessung in der Entwicklungszusammenarbeit durch die konzeptionelle Neuausrichtung und Veränderung entwicklungspolitischer Verfahren und Instrumente noch mehr gesteigert. Für einige Instrumente (Länderprogramme, Sektorprogramme, programmorientierte Gemeinschaftsfinanzierung, Budgethilfen) kann bislang nur auf wenige Ansätze und auf noch nicht völlig ausgereifte Verfahren zur Analyse ihrer Wirksamkeit zurückgegriffen werden. Wie bereits in HUMMELBRUNNER et al. in diesem Band aufgezeigt, bilden dynamische und kontextbezogene Betrachtungen von Wirkungen, welche der Komplexität und den Besonderheiten sozialer Systeme gerecht werden, bis dato noch Ausnahmen (vgl. AEG SAE 2013: 2, 23 ff.). Überdies bringen es die tendenzielle Zunahme der AkteurInnen in der Entwicklungszusammenarbeit und die Komplexität der Programme und Maßnahmen mit sich, dass zur Wirkungsmessung noch mehr Daten erforderlich werden, welche von einer wachsenden Anzahl an Organisationen geliefert werden müssen. Allein die dafür notwendige Koordination und Kooperationsbereitschaft der AkteurInnen gestaltet sich in der Praxis oftmals als problematisch (vgl. ARBEITSKREIS EVALUATION VON ENTWICKLUNGSPOLITIK DER DEGEVAL 2009: 11). Zusätzlich kann die durch Maßnahmen und Programme verfolgte Vielfalt an anvisierten Zielen im Kontrast zu einer robusten Wirkungsanalyse stehen, die von Beginn an die Maßnahmen und Programme auf bestimmte Wirkungen ausrichtet und damit eingrenzt.

Eine weitere enorme Herausforderung manifestiert sich in dem hohen Unsicherheitsgrad der Rahmenbedingungen in denen Entwicklungszusammenarbeit umgesetzt wird:

„Die Zielsysteme in denen die Maßnahmen umgesetzt werden sollen – und dies kann zweifelslos insbesondere für den Bereich der Entwicklungszusammenarbeit bestätigt werden – zeichnen sich oftmals durch ein hohes Maß an Unsicherheit, beispielsweise im Hinblick auf die politischen, wirtschaftlichen, sozialen, ökologischen (etc.) Rahmenbedingungen, Akteurskonstellationen oder Entwicklungsdynamiken, aus. Programmplaner agieren dementsprechend in hoch volatilen Umwelten, deren Zukünfte kaum abschätzbar und schon gar nicht prognostizierbar sind“ (SILVESTRINI 2011: 164).

DISKREPANZEN IN DER WIRKUNGSBEURTEILUNG

Den in den letzten Jahrzehnten durchgeführten internen und externen Evaluationen zur Wirksamkeit der Entwicklungszusammenarbeit Glaube geschenkt, zeigten sich auf der Projekt- und Programmebene (Mikroebene) durchweg

positive Evaluationsergebnisse. Mehr als drei Viertel der durchgeführten Projekt- und Programminterventionen wurden als zufriedenstellend und besser bewertet. Jedoch wurden auf der Ebene der Gesamtgesellschaft (Makroebene) über statistische Verfahren des Ländervergleichs bestenfalls widersprüchliche Befunde zum Zusammenhang zwischen Entwicklungstransfers und Entwicklung berichtet (vgl. STOCKMANN et al. 2010: 472; FAUST 2010). Dieses als Mikro-Makro-Paradoxon deklarierte Phänomen lässt sich folgendermaßen herleiten:

Die Makroebene betreffend, wird argumentiert, dass mit der Abhängigkeit einer Volkswirtschaft von den Transferleistungen der Entwicklungszusammenarbeit negative gesamtwirtschaftliche Effekte einhergehen können, welche die positiven Wirkungen gelungener Einzelprojekte kompensieren. Auf der Mikroebene wird häufig das Paradoxon in der Weise aufgelöst, indem an den positiven Ergebnissen der Projektevaluierungen gezweifelt wird. Als Argument wird die unzureichende empirische Evidenz der Evaluationsergebnisse in den ersten Dekaden der Entwicklungszusammenarbeit hervorgehoben. Die organisatorisch-institutionelle Kritik macht hierbei relevant, dass die Evaluationsschwerpunkte in der Vergangenheit zu sehr auf Input und Output von Maßnahmen gelegt wurden und die Ebene der Impacts überwiegend im Hintergrund verblieb. Demgemäß reicht alleine der Blick auf die Quantität der eingesetzten Ressourcen bzw. deren unmittelbare Ausgestaltung in Projekten nicht aus, um Informationen zu den Wirkungen auf der Ebene der Zielgruppe zu gewinnen. Weiterhin werden die isolierten Evaluationsaktivitäten der Geberorganisationen sowie die eingeschränkte Unabhängigkeit von Evaluation moniert, die zur Verzerrung der Ergebnisse zugunsten der Organisationen beitragen. Die inhaltlich-methodische Kritik rekurriert auf die Art und Weise, wie Wirkungsmessungen umgesetzt werden. Sowohl die häufig mangelhafte Datenerhebung als auch das Fehlen von Baseline-Studien zu Projektbeginn als Grundlage für Wirkungsmessungen bilden hier den Fokus (vgl. FAUST 2010).

Diskrepanzen in der Wirkungsbeurteilung erwachsen neben methodischen und institutionellen Schwachstellen, insbesondere auch aus der grundlegenden begrenzten Aussagekraft von Daten. NUSCHELER (2008: 25) hebt in diesem Zusammenhang die aus quantitativen Makrodaten abgeleiteten Schlussfolgerungen hervor, die qualitativ interpretiert werden müssen und viele Entwicklungen auf lokaler, gesellschaftlicher und politischer Ebene nicht abbilden können. Was bei NUSCHELER nur unterschwellig und in der deutschen Fachliteratur marginal zur Sprache kommt, tangiert das darüber hinausgehende generelle Interpretationsproblem von Daten (vgl. WILHELM 2012: 59 f.). In Großbritannien hingegen ist die Debatte schon etwas weiter. So konstatiert Peel betreffend Multi-Stakeholder-Evaluation Folgendes: „Different stakeholders will construe evaluation methods and filter the data differently, motivated by alternative value stances and priorities, information and learning needs.“ (PEEL 2003: 159)

2.4 Flexibilität von Planung und Zielverfolgung

Wie bereits angekündigt, zeichnen sich die Zielsysteme, in denen die Maßnahmen der Entwicklungszusammenarbeit umgesetzt werden sollen durch ein hohes Maß an Unsicherheit im Hinblick auf politische, wirtschaftliche, soziale und ökonomische Rahmenbedingungen aus (vgl. SILVESTRINI 2011: 164). Dies macht unter anderem eine relativ kurze Halbwertszeit abgestimmter Planung erforderlich (vgl. WILHELM 2012: 65), um flexibel auf unerwartete Entwicklungen reagieren zu können. Aus dieser Warte betrachtet, erscheint es wenig sinnvoll, bereits zu Beginn eines Programms mit unveränderbaren Zielformulierungen zu operieren, denn Ziele werden sich im Laufe der Umsetzungsphase noch verändern. Jedoch enthält nahezu jeder Evaluationsauftrag die Aufgabe zu überprüfen, inwiefern die zu Beginn eines Programms angestrebten Ziele erreichbar sind (formativ) oder erreicht wurden (summativ). Demnach bilden die Identifizierung von Programmzielen und die Beurteilung der Zielerreichung wesentliche Bestandteile von Evaluationen, welche allerdings in ihrer Umsetzung mit großen Herausforderungen verbunden sind (vgl. STOCKMANN 2007: 64 f.). Als Grundlage dienen Indikatoren, die entsprechend der Zielsetzungen der Programme und Projekte festgelegt werden und einen Soll-Ist-Vergleich ermöglichen sollen. Hierzu werden die Indikatoren in ihren Ausprägungen, so weit möglich, von Beginn der Implementation einer Intervention bis zum Abschluss mittels Monitoringsystemen beobachtet (vgl. MEYER 2004: 3). Damit kann einer Weiterentwicklung oder Änderung von verfolgten Zielen, wie dies häufig in der Praxis der Fall ist, folglich kein Augenmerk geschenkt werden.

2.5 Divergierende Erwartungen an Evaluation

Die Funktionen, die der Evaluation im Allgemeinen zugeschrieben werden, oszillieren zwischen verschiedenen, teils sogar gegensätzlichen, Ausprägungen. Demnach kann Evaluation zur Legitimation gegenüber dem Auftraggeber oder der Gesellschaft dienen, sowie unter anderem zur Qualitätssicherung über die Generierung von Erkenntnissen. Die in diesem Beitrag bislang noch nicht beleuchteten, aber mindestens ebenso bedeutsamen Funktionen, bilden die der Kontrolle sowie die des Lernens (vgl. STOCKMANN 2007: 36). Hierbei wird häufig in der Evaluationsforschung relevant gemacht, dass diese beiden Funktionen miteinander konfliktieren. So sollen einerseits Erkenntnisse gewonnen werden, die einen Beitrag zur Qualitätsweiterentwicklung und -sicherung leisten und somit auch das Lernen aus Fehlern implizieren, während andererseits die Funktion der Kontrolle, teils mit finanziellen Sanktionen verbunden,

eben genau das erste beeinträchtigen oder sogar verhindern kann (vgl. KÜHL 2008: 65 ff.; FROSCHAUER 2003).

Diese Kategorisierung der Evaluationsfunktionen soll jedoch nicht zu der Annahme verleiten, dass eine konkrete Evaluation z. B. im Rahmen einer Programmevaluierung in der Entwicklungszusammenarbeit eindeutig einer Funktion zugeordnet werden könnte. Vielmehr sind „sehr unterschiedliche Funktionen mit ein und derselben Evaluation auch gleichzeitig innerhalb einer Organisation denkbar“ (WILHELM 2012: 236). Welche Funktionen einer Evaluation zugeschrieben werden, korrespondiert zum einen mit autonomen Entscheidungen, welche die zu evaluierende Organisation selbst trifft und steht zum anderen in Abhängigkeit vom Organisationskontext. Je nachdem welche Funktion(en) eine Evaluation erfüllen soll, entscheidet letztlich auch über den Umgang mit den Evaluationsergebnissen (vgl. WILHELM 2012: 229 ff.).

Ausgehend von diesen Überlegungen sei zu berücksichtigen, dass die Wahrscheinlichkeit des Auftretens widersprüchlicher Erwartungen mit der Anzahl der an einer Evaluation Beteiligten steigt (vgl. PEEL 2003: 160). Wie mehrfach zum Ausdruck gebracht wurde, zeichnet sich die Entwicklungszusammenarbeit ohnehin durch eine Vielzahl beteiligter AkteurInnen aus, doch lässt sich in den letzten Jahren noch eine Multiplizierung der AkteurInnen in der Entwicklungszusammenarbeit auf regionaler, nationaler und globaler Ebene identifizieren. Die Diversität der AkteurInnen-Landschaft wird durch folgende Punkte evoziert:

- a) Auf der globalen Ebene findet eine grundlegende Verschiebung der Kräfteverhältnisse statt.
- b) Die Multiplizierung der Ziele der Entwicklungszusammenarbeit bringt auch neue AkteurInnen mit sich.
- c) Die Zunahme des öffentlichen Bewusstseins für globale Abhängigkeiten bedingt ebenso eine wachsende Zahl an AkteurInnen.
- d) Die Ausweitung der Ziele macht eine stärkere Unterteilung der Hilfen in Themen und Subthemen erforderlich, die neuen Organisationen die Möglichkeit bietet, sich in den Markt zu etablieren.
- e) Neue Organisationen ersetzen traditionelle, da diese in ihrer Leistung nicht überzeugt haben (vgl. STOCKMANN, BORRMANN 2009a: 38).

Im Lichte dieser Erkenntnisse sollte es daher Anspruch einer Evaluation sein, den unterschiedlichen Erwartungen der Beteiligten Rechnung zu tragen, um zu guter Letzt den Nutzen, die Glaubwürdigkeit und die Akzeptanz von Evaluationsergebnissen zu erhöhen. Andernfalls könnte eine Verdrossenheit gegenüber Evaluationen bei den Beteiligten die Folge sein (vgl. WILHELM 2012: 67 f.).

3 ZWISCHENFAZIT

Wie die bisherigen Ausführungen deutlich machen, ist die Entwicklungszusammenarbeit von einem enormen Legitimationsdruck geprägt. Das Erzielen sichtbarer und vor allem nachweisbarer Erfolge steht nicht zuletzt auch aufgrund der Millennium Development Goals (MDGs) sowie der Paris-Deklaration besonders im Mittelpunkt. Evaluation erfährt in der Entwicklungszusammenarbeit eine große Bedeutung, um vor allem die Wirksamkeit zu belegen.

Um den großen Erwartungen Rechnung zu tragen, muss sich Evaluation einer nahezu unüberschaubaren Anzahl an Herausforderungen stellen, wie die anschließende Abbildung verdeutlicht:



Abb. 1: Herausforderungen von Evaluationen in der Entwicklungszusammenarbeit; Quelle: Eigene Darstellung

Diese ergeben sich insbesondere aus den entwicklungspolitischen Spezifika der Maßnahmen und Programme und den veränderten Umsetzungsmodalitäten. Hier sei der hohe Unsicherheitsgrad der Rahmenbedingungen, unter denen Entwicklungszusammenarbeit umgesetzt wird, benannt. Dieser führt dazu, dass Programmplanung und Zielformulierungen möglichst flexibel gehandhabt werden müssen, um auf unerwartete Entwicklungen reagieren zu können. Hinzu kommt die strategische Neuausrichtung der Entwicklungszusammenarbeit, die zunehmend auch auf multidimensionale Herangehensweisen fokussiert und eine wachsen-

de Zahl an AkteurInnen mit sich bringt. All dies mündet in einem drastischen Anstieg der Komplexität von Maßnahmen in der Entwicklungszusammenarbeit.

Vor dem Hintergrund dieser teilweise neuen Herausforderungen wird in der Entwicklungszusammenarbeit fortwährend die Frage nach adäquaten Evaluationsmodellen und -ansätzen in den Raum gestellt. Dabei haben aktuelle Diskussionen um Komplexität den Blick auf systemische Evaluationsansätze gelenkt. Da in der Entwicklungszusammenarbeit bereits seit Jahrzehnten systemische Organisationsberatung und seit geraumer Zeit systemische Managementansätze verankert sind, wächst nun auch das Interesse an systemischen Evaluationsansätzen. Doch was wird unter systemische Evaluationsansätze verstanden und vor allem, welche Chancen bieten sie? Welchen Mehrwert besitzen systemische Evaluationsansätze gegenüber anderen? Diesen Fragen wird sich in den anschließenden Kapiteln eingehend gewidmet, da angesichts der beschriebenen Entwicklungen in der Entwicklungszusammenarbeit angenommen wird, dass systemische Evaluationsansätze viele ungeahnte Potenziale in sich vereinen. Hier könnte sich eine Lücke schließen, die die bislang in der Evaluation angewendeten Methoden und Ansätze nicht zu schließen vermögen. Gleichermäßen wird sich der Frage zugewendet, welche Grenzen erreicht werden (können) und welche Hindernisse es zu überwinden gilt, wenn systemische Ansätze in der Evaluation zur Anwendung kommen.

4 BEDEUTUNGSAUFLADUNG SYSTEMISCHER EVALUATIONSANSÄTZE UND IHRE ANWENDUNG IN DER PRAXIS

Dieses Kapitel möchte Verständnisse und Charakterisierungen von systemischen Evaluationsansätzen in der deutschen staatlichen Entwicklungszusammenarbeit näher beleuchten. Hierzu wurden ExpertInnen befragt, welche sich entweder selbst als systemische EvaluatorInnen bezeichnen oder ihrer Ansicht nach systemische Evaluationsansätze anwenden oder sich bereits ausführlich mit dem Thema befasst haben (z. B. über eine Mitwirkung an Workshops zu entsprechenden Themen). Die qualitative, nicht-standardisierte Erhebungsmethode der Leitfadeninterviews ermöglichte dabei die Sichtbarmachung individueller Verständnisse, Erfahrungen und Bewertungen der InterviewpartnerInnen bezüglich systemischer Evaluationsansätze. Betreffend der Auswertung der Interviews kam die Methode der qualitativen Inhaltsanalyse nach MAYRING zur Anwendung (vgl. LAMNEK 2005: 517 ff.) und darüber hinaus ein hermeneutisches Vorgehen (Hermeneutischer Zirkel I und II) (vgl. ebd.: 61 ff.). Es wurde sich aus dem Grund zusätzlich für die Hermeneutik entschieden, da versucht wurde, „den subjektiv gemeinten Sinn“ (GIESEN, SCHMID 1976: 165 zit. n. LAMNEK 2005: 62) und damit das individuelle Verständnis der InterviewpartnerInnen bezüglich systemischer Evaluationsansätze zu identifizieren. Damit stand nicht allein das Gesagte der InterviewpartnerInnen im Fokus, sondern die tieferliegenden Annahmen, Wertmaßstäbe und Sinnstrukturen.

4.1 Erste Annäherung – Was wird unter systemischen Evaluationsansätzen verstanden?

In Bezug auf die Charakterisierung systemischer Evaluationsansätze und der Abgrenzung zu anderen Herangehensweisen kristallisiert sich heraus, dass bislang keine Definitionshoheit existiert. Die InterviewpartnerInnen schreiben daher systemischen Evaluationsansätzen teils ähnliche, aber teils auch sehr unterschiedliche Bedeutungen zu.

ALLGEMEINER KONSENS

Es herrscht ein allgemeiner Konsens darüber, dass sich systemische Evaluationsansätze insbesondere für den Umgang mit komplexen Situationen und Evaluationsgegenständen eignen. Hierbei werden durchgehend die Phänomene der Komplexität, Unvorhersehbarkeit, Nicht-Linearität, Mehrperspektivität und Reflexivität genauer zur Sprache gebracht, die speziell für den Bereich der Entwicklungszusammenarbeit prägend sind. Bedingt durch die Zusammenarbeit unterschiedlicher Akteure und dem Aufeinandertreffen verschiedener Interessenskonstellationen besitzt die Entwicklungszusammenarbeit ohnehin einen immanenten systemischen Kern. In diesem Zusammenhang wird in einem Interview folgende Erkenntnis formuliert:

„Wenn man nicht ein gewisses systemisches Grundverständnis hat, kann man hier gar nicht arbeiten“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Dennoch kommt durch die Interviews durchgehend ans Licht, dass, jenseits eines systemischen Grundverständnisses, die Art und Weise sowie auch die Intensität systemischen Arbeitens innerhalb der Entwicklungszusammenarbeit mannigfaltig ausfallen.

Darüber hinaus teilen, bis auf eine Ausnahme, alle InterviewpartnerInnen das Verständnis, dass Evaluation als Intervention zu fassen ist. Demnach wird sich von der Erwartungshaltung distanziert, Sachverhalte objektiv und total unabhängig festhalten und bewerten zu können:

„Es gibt viele Evaluatoren, die denken, ich gehe irgendwo rein und guck mir das von außen an. Anschließend gehe ich wieder raus und beschreibe es so wie es ist. In einer systemischen Evaluierung geht man von folgendem aus: ‚In dem Augenblick, wo ich meinen Fuß in die Organisation setze, fängt eine Intervention an‘“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Weitere Gemeinsamkeiten finden sich in der Abgrenzung systemischer Ansätze zu anderen Evaluationsansätzen. So werden rigorose Evaluationsansätze, Soll-Ist-Vergleiche, Nutzen-Kosten-Analysen als Gegenbeispiele zur Sprache gebracht, um systemische Evaluationsansätze zu charakterisieren.

Insgesamt nähern sich damit alle InterviewpartnerInnen einer Begriffsklärung an und dementieren jedoch eine klare Abgrenzbarkeit und Definierbarkeit.

DIVERGIERENDE AKZENTUIERUNGEN

Zwar thematisieren alle InterviewpartnerInnen die Phänomene der Komplexität, Unvorhersehbarkeit, Nicht-Linearität, Mehrperspektivität und Reflexivität, um systemische Evaluationsansätze greifbar zu machen, dennoch fallen die Akzentuierungen entlang dieser Komponenten jeweils divergent aus. In einem Interview werden bspw. besonders die Komplexität und Unvorhersehbarkeit von Entwicklungen und Vorhaben in der Entwicklungszusammenarbeit sowie die Nicht-Linearität von Kausalzusammenhängen in den Mittelpunkt gerückt, während auf den Bereich der Reflexivität nur implizit eingegangen wird. In einem anderen Gespräch wird hingegen die Reflexivität als Kernmerkmal systemischen Arbeitens herausgestellt und wiederum in einem weiteren Interview rücken vorrangig die Phänomene der Komplexität und Mehrperspektivität ins Blickfeld. Die anderen Komponenten finden zwar Erwähnung, fließen jedoch mit einem deutlich geringeren Gewicht in die Gespräche mit ein. Aufgrund der fehlenden Definitionshöhe entwerfen die InterviewpartnerInnen daher eigene Begriffsannäherungen, die, wie die folgende zeigt, systemische Evaluationsansätze lediglich konturenhaft beschreiben:

„Es gibt unterschiedliche Wahrnehmungen von betroffenen Akteuren, die alle gleichermaßen ernst zu nehmen sind. Darüber hinaus geht man prozessorientierter in der Evaluierung vor. Man bezieht mehr Faktoren ein, die auch jenseits von Indikatoren wichtig sind. Die berücksichtigt man, auch wenn das in den ‚terms of reference‘ nicht direkt steht, denn es könnte sein, dass das relevant ist. Man versucht mehr das Gesamte zu erfassen. Das versucht man in anderen Evaluierungen auch, aber es steht dann doch im Fokus, nur bestimmte Indikatoren oder Fragestellungen im engeren Sinne zu bedienen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Überdies kommen die unterschiedlichen Verständnisse von systemischen Evaluationsansätzen in einer uneinheitlichen Begriffsverwendung zum Vorschein. So sprechen einige InterviewpartnerInnen ausdrücklich von ‚systemischen Evaluationsansätzen‘ und andere von ‚systemischer Evaluation‘. In diesem Kontext entgegnet eine Gesprächsperson:

„Ich bin kein systemischer Evaluator. Ich verwende systemische Ansätze in der Evaluation“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Die Bewerkstelligung einer systemischen Evaluation wird in diesem Interview jedoch nicht bestritten. Im Kontrast dazu hebt eine andere Gesprächsperson hervor, dass es niemals eine systemische Evaluation geben kann:

„Es gibt keine systemische Evaluation, sondern es gibt nur systemische Ansätze in der Evaluation. Das ist ganz wichtig. Systemische Evaluation, würde ich sagen, kann es nicht geben“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

In einem weiteren Gespräch wird zu der Frage nach der genauen Begriffswahl Folgendes geantwortet:

„Mir geht es darum, möglichst viel systemische Sichtweise in Entwicklungsprozesse zu bekommen, egal ob es die Evaluierung, die Beratung oder die Planung betrifft. Über die terminologischen Feinheiten mache ich mir eigentlich keine Gedanken“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Auch wenn einige InterviewpartnerInnen übereinstimmend von systemischen Evaluationsansätzen sprechen, scheinen diese dennoch Unterschiedliches unter diesem Begriff zu subsumieren. Einerseits werden unter systemischen Evaluationsansätzen konkrete Instrumente oder auch Methoden gefasst, die in einer Evaluation zur Anwendung kommen können und angepasst an den jeweiligen Evaluationsgegenstand situationsspezifisch ausgewählt werden und andererseits werden damit über Methoden hinausgehende Herangehensweisen bezeichnet. Hinzu kommt, dass die Begriffswahl auf der Grundlage unterschiedlicher Begründungen und Argumente erfolgt. So verweist eine Gesprächsperson darauf, dass nicht von einer systemischen Evaluation die Rede sein kann, da es keine klare Abgrenzbarkeit zu nicht-systemischen Evaluationen gibt. Abweichend dazu wird in einem anderen Interview das Argument untermauert, dass lediglich systemische Methoden und Instrumente zur Anwendung kommen und sich dahinter eine wissenschaftstheoretische Frage verbirgt.

Korrelierend mit den unterschiedlichen Verständnissen von systemischen Evaluationsansätzen wird auch die Reichweite dieser Herangehensweisen nicht einheitlich abgeschätzt wie die folgende Abbildung illustriert:



Abb. 2: Divergierende Verständnisse von systemischen Evaluationsansätzen; Quelle: Eigene Darstellung

Demgemäß kristallisiert sich in einem Interview heraus, dass, aus einer systemisch-konstruktivistischen Brille betrachtet, nicht nur die bisher in der Evaluationspraxis gängigen Methoden und Verfahren überdacht werden müssen, sondern das gesamte Evaluationssystem:

„Mein persönliches Verständnis wäre, dass man in einer systemischen/konstruktivistischen Denkweise das gesamte System der Evaluierung eigentlich überdenken müsste. Also was heißt das überhaupt? Machen dann ex-post-Evaluierungen Sinn und wie machen die Sinn?“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Auch der Stellenwert von Indikatoren wird in diesem Zusammenhang hinterfragt. Demgegenüber fassen andere InterviewpartnerInnen die Reichweite wesentlich enger, betonen jedoch eine spezifische Haltung, die für die Anwendung systemischer Ansätze von fundamentaler Bedeutung ist. Diese spezifische systemische Haltung ist zwar mit Implikationen für das Evaluationssystem verknüpft, allerdings wird in den Interviews inhaltlich nicht näher darauf eingegangen. Inwieweit systemische Evaluationsansätze über ihre Anwendung durch die EvaluatorInnen hinweg mit z. B. Indikatoren oder Evaluationszeitpunkten in Verbindung stehen, bleibt daher offen und spielt folglich für einige Gesprächspersonen eine eher untergeordnete Rolle.

ROLLE DER EVALUATORINNEN: „OBJEKTIV GEHT NICHT“

Im Hinblick auf die Rolle und Funktion von EvaluatorInnen teilen die InterviewpartnerInnen ein sehr ähnliches Bild. Demzufolge nehmen systemisch inspirierte EvaluatorInnen unterschiedliche Rollen im Laufe des Evaluationsprozesses ein, im Sinne z. B. von ModeratorInnen, BeraterInnen, ImpulsgeberInnen und Coaches. Dem wird die Vorstellung gegenübergestellt, dass EvaluatorInnen als PrüferInnen oder RichterInnen objektives Recht sprechen sollen. In diesem Zusammenhang wird darüber hinaus der stark intervenierende Charakter von Evaluation immer wieder aufgezeigt:

„Ich persönlich bin der Meinung, sobald ich mich in den Dialog gebe, stoße ich das System schon an und bringe es aus dem Gleichgewicht. Ich bringe einen Prozess in Gang, der Wirkungen und eine Dynamik im System erzeugt. Das heißt, ein Evaluator begibt sich zwangsläufig in das System und ist in irgendeiner Form Teil des Systems“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Diese theoretischen Vorüberlegungen wirken sich unausweichlich auf die Rolle und das Selbstverständnis von EvaluatorInnen aus und führen unter anderem eine Vorgehensweise in Gestalt zirkulärer Schleifen mit sich. Inwiefern sich dieser Grundgedanke, dass EvaluatorInnen in ein für sie fremdes System intervenieren, in der Arbeitsweise niederschlägt, expliziert das folgende Zitat:

„Ich kann nicht so tun als sei ich ein neutrales Etwas. Schon durch die Tatsache, dass ich durch jemand Bestimmtes geschickt werde. Ich muss mir schon vorher überlegen, wie das wahrgenommen wird und wie ich dort auftrete. Ich muss also auch da schon das System betrachten. Was erzählen sie mir möglicherweise, wenn ich als Evaluator im Auftrag von jemandem komme und was erzählen sie mir genau deshalb nicht! Was würden sie mir vielleicht erzählen, wenn ich von jemand anderem käme?“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Hinsichtlich der Generierung von Evaluationsergebnissen wird die Beziehung zwischen EvaluatorInnen und Evaluierten thematisiert. Einmütig machen die InterviewpartnerInnen darauf aufmerksam, dass es als nicht zielführend angesehen wird, Evaluationsergebnisse ‚im stillen Kämmerlein‘ und damit unter Ausschluss der Evaluierten zu erstellen. Diese Erkenntnis sei jedoch nichts Spezifisches für systemische Evaluationsansätze, sondern charakterisiere vielmehr den generellen Trend der letzten Jahrzehnte in der Evaluationspraxis. Inwiefern die Evaluierten in der Praxis tatsächlich in den Evaluationsprozess involviert sind und damit an der Generierung von Evaluationsergebnissen beteiligt, geht aus den Interviews häufig nicht klar hervor. Allerdings bildet eine Gesprächsperson eine Ausnahme

und rückt den stark partizipativen und reflexiven Charakter systemischen Arbeitens in den Mittelpunkt:

*„In einer systemisch angelegten Evaluierung wird man immer wieder spiegeln. Das heißt, es gibt bestimmte Teilergebnisse, die man abcheckt mit anderen, durch Triangulation. So kann man gucken, ob man das verstanden hat oder ob sich in Kommunikation mit anderen Akteuren noch etwas Neues ergibt. Das heißt, die Organisation ist am Inhalt mit dabei. Wir haben es nicht mit einer Blackbox zu tun, wo am Ende ein Bericht erscheint und alle sind überrascht. Es gibt mit relevanten Akteuren einen Diskussionsprozess, der mehr oder weniger intensiv sein kann“ (INTERVIEW MIT EINEM EVALUATOR/
EINER EVALUATORIN).*

Kontrastierend zu dieser Perspektive wird der Themenbereich in anderen Interviews nur peripher ins Gespräch gebracht trotz gezielter Nachfragen. So exponiert eine Gesprächsperson ausdrücklich, dass Empfehlungen am Schluss einer Evaluation unerlässlich sind und zum Standard einer Evaluation gehören. Im Idealfall sollten diese Empfehlungen partizipativ und in einem Dialogprozess erarbeitet werden, doch ist sich diese Person nicht sicher, ob das in der Praxis auch tatsächlich so geschieht. Daraus kann abgeleitet werden, dass die Rollen und Arbeitsweisen systemischer EvaluatorenInnen in der Praxis von den InterviewpartnerInnen nur konturenhaft beschrieben werden können.

TRENNSCHÄRFE SYSTEMISCHER EVALUATIONSANSÄTZE – ZWISCHEN SCHWARZ-WEISS-GRENZE UND KONTINUUM

Grundsätzlich teilen die InterviewpartnerInnen zwar ein einheitliches Verständnis darüber, dass die Grenze zwischen systemischen und nicht-systemischen Herangehensweisen nicht trennscharf ist, dennoch sehen einige mehr als andere markante Unterschiede zwischen der Anwendung systemischer Evaluationsansätze und gegenwärtigen Evaluationsdesigns. In einem Gespräch wird in Verbindung mit einer geringen Trennschärfe vor allem die unzureichende Konzeptionalisierung von systemischen Evaluationsansätzen diskutiert. Diese Gesprächsperson zieht daher auch den Schluss, dass es bis dato kein geteiltes Wissen darüber gibt, wie man systemische Evaluationsansätze in der Praxis anwendet. Andere InterviewpartnerInnen hingegen verweisen auf für sie klare systemische Konzepte und Methoden. Dass diese systemischen Konzepte für einige jedoch durchaus Fragen aufwerfen, versinnbildlicht dieses Zitat:

„Was heißt denn das eigentlich? Also wir finden die Konzepte schon überzeugend, aber was es dann tatsächlich heißt? Da kann man so viele verschiedene Sachen drunter verstehen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Für viele InterviewpartnerInnen scheint darüber hinaus die Abgrenzung zu partizipativen Ansätzen zu verschwimmen. Grundlegend werden viele Gemeinsamkeiten zwischen systemischen und partizipativen Evaluationsansätzen identifiziert, allerdings scheiden sich die Geister hinsichtlich der Bewertung der geringen Unterscheidbarkeit. Einerseits wird dies als unproblematisch erachtet und andererseits wird gewünscht, dass sich die Wissenschaft dieser Fragestellung eingehender zuwendet, da ansonsten die Gefahr besteht, dass beide Ansätze fälschlicherweise miteinander gleichgesetzt werden. Im Hinblick auf die Frage, worin die Unterschiede zwischen systemischen und partizipativen Evaluationsansätzen liegen, repliziert eine Gesprächsperson Folgendes: „Man kann auch im Partizipativen linear denken und vorgehen“ (Interview mit einem Evaluator/einer Evaluatorin). Hiermit wird illustriert, dass partizipative Vorgehensweisen nicht automatisch einem systemischen Weltbild zugrundeliegen, aber umgekehrt integrieren systemische Arbeitsweisen einen partizipativen Kern.

Des Weiteren vertreten alle InterviewpartnerInnen die Meinung, dass die Grenzen zwischen Beratung und Evaluation zu verschwimmen scheinen, wenn systemische Evaluationsansätze zur Anwendung kommen. Diese Grenzverwischung wird dabei von einigen als unproblematisch bewertet, während andere an dieser Stelle einen konzeptionellen Handlungsbedarf erkennen:

„Wobei es auch Ansätze gibt, wie developmental evaluation, die auch als Beratung angelegt sind und ich finde da ist auch die Grenze zwischen Evaluierung und Beratung an der Stelle nicht mehr trennscharf. Ich bin auch nicht so überzeugt, ob das so gut ist. Ich finde, das sollte man schärfen, sonst ist nachher wieder alles Beratung und dann wird es irgendwie auch gummimäßig“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Evaluation immer mehr als Beratung zu verstehen, ist auch dem Umstand geschuldet, dass ein Großteil der methodischen Verfahrensweisen direkt aus der Organisationsberatung auf den Bereich der Evaluation übertragen wird. Auch hier teilen sich die Meinungen, ob eine derartige Transformation ohne weitere Einschränkungen und Anpassungen möglich ist. Eine Gesprächsperson sieht diese Vorgehensweise dabei besonders kritisch und konstatiert einen eher hemdsärmlichen Umgang mit diesem Thema.

Ein weiterer interessanter Aspekt erwächst aus der Annahme, dass die mangelnde Trennschärfe systemischer Vorgehensweisen ein der Entwicklungszusammenarbeit inhärentes Phänomen darstellt:

„In der Entwicklungszusammenarbeit arbeiten viele Systemiker, so dass ich glaube, dass die Abgrenzung innerhalb der Entwicklungszusammenarbeit, in dem was das systemische Arbeiten und nicht-systemische Arbeiten angeht, schwer fällt. Da ist wahrscheinlich eine Abgrenzung zu anderen Branchen leichter“
(INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Gemäß dieser Gesprächsperson wäre es auch denkbar, sich die Trennschärfe systemischer Evaluationsansätze weniger als eine Schwarz-Weiß-Grenze, sondern als ein Kontinuum vorzustellen. Eine Einteilung in ‚systemisch‘ oder ‚nicht-systemisch‘ wird dadurch obsolet, aber auf der Basis eines Vergleiches und eines gegenseitigen Abwägens von Evaluationsaktivitäten und ihrer systemischen Komponenten wäre eine relationale Zuordnung in Richtung ‚mehr systemisch‘ oder ‚weniger systemisch‘ erwägenswert.



Abb. 3: Wie neu sind systemische Evaluationsansätze? Quelle: eigene Darstellung

4.2 Systemische Evaluationsansätze in der Praxis

„ALTER WEIN“ ODER NEULAND?

Den Neuheitsgrad systemischer Evaluationsansätze betreffend, gehen die Meinungen weit auseinander. Einen Überblick über die von den InterviewpartnerInnen vertretenen Einschätzungen liefert Abbildung 3.

Einige konstatieren, dass in der Evaluationspraxis schon seit längerer Zeit systemische Methoden angewendet werden, diese aber nicht als solche explizit benannt wurden. Lediglich der Diskurs um systemische Vorgehensweisen in der Evaluation sei relativ neu, wie das anschließende Zitat kommuniziert:

„Wenn man annimmt, dass seit ungefähr 50 Jahren Evaluierungen durchgeführt werden, dann ist das auf die Perspektive von 50 Jahren gesehen eine etwas neuere Entwicklung. Aber eine, die mindestens in den 90er Jahren eingesetzt hat, wahrscheinlich schon früher. Nur jetzt wird es auch öffentlich diskutiert z. B. in der DeGEval und an anderen Stellen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Diejenigen, die diese Ansicht vertreten, waren oder sind noch immer schwerpunktmäßig auf dem Gebiet der Organisationsberatung und -entwicklung tätig, in dem das systemische Denken bereits vor Jahrzehnten Einzug gehalten hat. Dass nicht alle InterviewpartnerInnen mit dieser Ansicht konform gehen, verdeutlicht das folgende Zitat:

„Die systemischen Ansätze per se sind ja nicht so neu. Das ist eine Methodentradition, die reicht schon 50, 60 Jahre zurück. Die Anwendung in der Evaluation ist ein relativ neues Phänomen und reicht, so wie ich das überblicke, ungefähr zehn Jahre zurück. Und hier nimmt das Interesse durchaus zu. Also mir geht es nicht rasch genug. Hier sollte man in der Anwendung schon wesentlich weiter sein“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Im starken Kontrast dazu wird in einem anderen Interview die Position vertreten, dass systemische Evaluationsansätze äußerst jung und sogar unausgereift wären. Es wird dementiert, dass systemische Herangehensweisen seit längerer Zeit in der Evaluation zur Anwendung kommen und hervorgehoben, dass sich systemisches Denken von gegenwärtigen Evaluationsdesigns deutlich absetzt. Der Neuheitsgrad wird als sehr hoch bewertet, doch würde das tatsächlich Neue an

systemischen Evaluationsansätzen aufgrund einer mangelnden Reife und bislang unzureichenden Konzeptionalisierung bis dato nur schwer fassbar.

„Meiner Meinung ist die Anwendung in der Praxis konzeptionell nicht weit genug ausgereift, deswegen ist es schwierig, gute Gutachter zu finden, die das auch anwenden können. Am Ende unserer Evaluierung, die viel Geld gekostet hat, stellten sich uns daher schon die Fragen ‚War daran jetzt irgendwas neu und was hat das jetzt eigentlich mehr gebracht?‘“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Eine weitere Gesprächsperson ist sich unschlüssig über den Neuheitsgrad und kann die Frage, ob systemische Evaluationsansätze in der Praxis nun eher neueren oder schon älteren Datums sind, nicht eindeutig beantworten. Dabei führt diese Person die konträren Positionen, die systemisches Arbeiten in der Evaluation entweder als älter oder als äußerst jung betrachten, zusammen:

„Ich denk der Begriff ist schon noch neuer. Ob das in der Vorgehensweise neu ist, das weiß ich selbst nicht genau. Vielleicht schafft man auch einfach nur neue Begriffe, um auch Dinge vermarkten zu können. Dann sprechen wir von einem Ansatz, der sinnvoll ist, der vielleicht nicht ganz neu ist. Aber es entstehen ja selten ganz neue Sachen, sondern immer Dinge aus etwas Bestehendem. Daraus wird dann was zusammengemixt, so wie bei einem Teig. Der entsteht ja nicht aus einem Nichts heraus, sondern man braucht bestimmte Zutaten, die schon da sind“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

In Anbetracht der Auffassung, dass systemische Evaluationsansätze nicht ganz neu sind und sich vielmehr aus etwas bereits Bestehendem entwickelt haben und noch entwickeln, wird indes nicht ausgeschlossen, dass systemische Evaluationsansätze neue Elemente enthalten, die sich von der aktuellen Evaluationspraxis abgrenzen. So entgegnet diese Person an einer anderen Stelle des Gesprächs, dass sich systemische Evaluationsansätze noch in einem Findungsprozess befinden und sich die Unterschiede zu gegenwärtigen Evaluationsdesigns eventuell erst im Laufe der Zeit stärker herauskristallisieren. Diese Einschätzung weist daher eine größere Schnittmenge mit der Position derjenigen Gesprächsperson auf, die die unzureichende Konzeptionalisierung im Kontext der Bewertung des Neuheitsgrades unterstreicht.

Gebührend der Metapher des Teigs, der nur mit bereits vorhandenen Zutaten gären kann, könnten diese Zutaten übertragen auf systemische Evaluationsansätze unter anderem partizipative Ansätze darstellen. Vor diesem Hintergrund gewinnt auch das Zitat einer weiteren Gesprächsperson an neuer Bedeutung:

„Zirkuläre Denkweise oder iterative Schleifen, die in den 80ern aufgekommen sind, sind klassisch systemisch, weil das vom linearen Denken weggeht. Das ist gerade bei den partizipativen Ansätzen, wo Jules Pretty und auch Chambers in England federführend waren, eingeführt worden. Man hat es damals einfach nicht systemisch genannt, aber das ist klassisch systemisch“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Ohne die Metapher des Teigs könnte das voranstehende Zitat so gedeutet werden, dass systemische Evaluationsansätze nicht mehr als ‚alter Wein in neuen Schläuchen‘ sind, stattdessen offenbart sich im Laufe des Gesprächs, dass diesen durchaus etwas ganz Neues inhärent ist:

„Das ganz explizit Systemische ist ganz neu und ich weiß nicht, ob man sagen kann, es gibt schon so viele Evaluationen, wo explizit systemische Ansätze genannt worden sind“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Die Einschätzungen der Interviewten in Bezug auf den Neuheitsgrad systemischer Evaluationsansätze nährt daher die Erkenntnis, dass es sich hier um ein relativ junges Phänomen handelt, dessen Vorläufer sich unter anderem in partizipativen Ansätzen manifestieren, die schon einige Jahrzehnte zurückreichen. Viele Grundannahmen und Vorgehensweisen bei systemischen Evaluationsansätzen speisen sich daher aus bereits gängigen Praktiken und inkludieren dennoch Elemente, die bislang in der Evaluationspraxis selten bis gar nicht in Erscheinung treten. Da die Konzeptionalisierung noch am Anfang steht und damit ebenso die Sichtbarmachung des Charakteristischen und Neuen an systemischen Evaluationsansätzen, gestaltet sich die Bewertung des Neuheitsgrades als äußerst kompliziert und bringt folglich unterschiedliche Positionen ans Licht.

BEWUSSTER UND UNBEWUSSTER EINSATZ SYSTEMISCHER EVALUATIONSANSÄTZE

Wie bereits erwähnt, gehen einige InterviewpartnerInnen davon aus, dass systemische Herangehensweisen schon seit geraumer Zeit das Vorgehen in der Evaluationspraxis prägen. Hierbei tritt in Augenschein, dass diejenigen EvaluatorenInnen, die auch Organisationsberatung und -entwicklung anbieten, ausbildungsbedingt ein systemisches Grundverständnis mitbringen. Bei ihnen finden systemische Elemente im Sinne eines ‚common sense‘ Zustimmung in der Anwendung und werden eher unbewusst eingesetzt bzw. nicht ausdrücklich als systemisch etikettiert. EvaluatorenInnen, die keine Ausbildung in der Organisationsberatung durchlaufen haben, sind auf dem Gebiet systemischer Ansätze oftmals weniger versiert:

„Es gibt nicht so viele, die sich als systemisch bezeichnen und die sich mit Systemtheorie befassen haben. Das sind dann eher Berater und die machen auch Organisationsentwicklung. Da gibt es viele. In der Evaluation, also die, die sich als reine Gutachter verstehen, haben sich häufig nicht mit systemischen Ansätzen befasst“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Insofern sagt die Gesprächsperson zwar ausdrücklich, dass systemische Ansätze in der Evaluation schon seit längerer Zeit angewendet werden, nimmt aber eine Einschränkung vor.

Darüber hinaus dementieren zwei InterviewpartnerInnen ganz klar eine etablierte Anwendung systemischer Ansätze in der Evaluationspraxis. Dabei ist den beiden gemeinsam, dass sie systemische Ansätze grundlegend affirmieren und eine verbreitete Anwendung dieser sehr begrüßen würden. Die bis dato geringe Berücksichtigung systemischer Ansätze wird daher ausdrücklich kritisiert und in Kontrast gestellt zu der Bereitschaft angelsächsischer Länder, in welchen systemische Evaluationsansätze schon seit längerer Zeit auf größere Resonanz und Aufmerksamkeit stoßen:

„Ich merke in der internationalen Diskussion mehr Bereitschaft, gerade in der angelsächsischen Welt. Hier wird die Notwendigkeit systemischer Ansätze schon weiter gesehen und es gibt auch mehr Bereitschaft, sich darauf einzulassen und das ganze eben zu erproben“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Die Gründe für die mangelnde Akzeptanz und Anwendung systemischer Herangehensweisen in der deutschen Entwicklungszusammenarbeit werden von den InterviewpartnerInnen allerdings unterschiedlich gesehen. Diesbezüglich wird relevant gemacht, dass bisher darin verharrt wird, mit herkömmlichen Instrumenten komplexen Evaluationsgegenständen zu begegnen und es schwerfällt, von diesem Weg abzuweichen. In einem anderen Interview wird der Grund in der mangelnden Konzeptionalisierung systemischer Evaluationsansätze gesehen sowie in dem fehlenden Rückhalt von starken BefürworterInnen, die systemischen Evaluationsansätzen ein Sprachrohr verleihen könnten:

„Die Debatte steht ja ganz am Anfang und bislang kann nicht klar herausgearbeitet werden, was daran eigentlich neu ist. Und so ist es wirklich schwer jemanden zu überzeugen, denn es gab kein richtig überzeugendes Argument und es gab niemanden, der es richtig gut vertreten konnte und da wirklich weit war“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

STEIGENDE AKZEPTANZ SEITENS DER AUFTRAGGEBERINNEN

Wie bereits angeklungen, sind sich die InterviewpartnerInnen nicht einig, inwieweit systemische Herangehensweisen bisher Einzug gehalten haben in die Evaluationspraxis. Jedoch wird durchgängig zur Sprache gebracht, dass die Debatte und das Interesse bezüglich systemischer Evaluationsansätze Konjunktur haben. Dies findet seinen Ausdruck darin, dass eine tendenziell wachsende Anzahl an Konferenzen organisiert wird, die sich explizit dem Thema der Komplexität widmen und damit systemischen Ansätzen zunehmend Gehör verschaffen. Diese Entwicklung markiert für einige einen großen Meilenstein, da die Vergangenheit ganz andere Reaktionen seitens der AkteurInnen in der Entwicklungszusammenarbeit zeichnete:

„Vor 2,5 Jahren hatten wir von einer ganzen Reihe an Wissenschaft, die rigoros impact-Fraktion vor allem, die rigoros und streng schematisch vorgehen, extremen Gegenwind. Da gab es sogar extrem kritische Bemerkungen aus dem Ministerium und wir hatten das Gefühl, wir müssen da richtig gegen eine Front kämpfen, die wenig bereit ist, die systemischen Ansätze zu akzeptieren. Inzwischen habe ich das Gefühl, dass sich der Wind gedreht hat und sie auch erkannt haben, dass es einen Mehrwert gibt“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Trotz des steigenden Interesses wird bislang noch sehr sensibel mit dem Thema umgegangen. So werden unter anderem Evaluationen, von EvaluatorsInnen ganz bewusst nicht als ‚systemisch‘ etikettiert, obwohl systemische Ansätze bewusst zur Anwendung kommen:

„Aber man kann das einfach oft nicht so verkaufen und das höre ich auch von welchen, die sich wirklich als systemische Evaluatoren verstehen. Die verkaufen das dann nicht als systemisch, sondern die verkaufen das als besonders geeignet für diesen Bereich“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Der sensible Umgang spricht für eine nach wie vor große Skepsis. Von zwei InterviewpartnerInnen wird sogar die Annahme geäußert, dass auf politischer Ebene andere Handlungslogiken vorherrschen, die mit systemischen Denkweisen eventuell gar nicht vereinbar sind.

Generell sei angemerkt, dass die Resonanz gegenüber systemischen Evaluationsansätzen von Seiten der AuftraggeberInnen von Evaluationsvorhaben nur schwer eingeschätzt werden kann. So erwähnt eine Gesprächsperson, dass oftmals diejenigen den Zuschlag für eine Ausschreibung erhalten, die deutlich weniger komplex denken. Allerdings könne lediglich gemutmaßt werden, woran dies im Detail liegt. Hierzu ergänzend thematisiert eine andere Person die Möglichkei-

ten und Fähigkeiten der AuftraggeberInnen, die Differenziertheit unterschiedlicher Evaluationsansätze samt ihrer Spezifika zu überblicken. Es wird davon ausgegangen, dass häufig schlicht und ergreifend die Ressourcen fehlen, um sich auf AuftraggeberInnen-Seite mit systemischen Evaluationsansätzen eingehend zu beschäftigen und diese dann als geeignete Variante zu identifizieren und zu affirmieren. Dass diesbezüglich Handlungs- und Aufklärungsbedarf besteht, wird in vielen Gesprächen kenntlich gemacht. Die direkte Erfahrungswelt der InterviewpartnerInnen, welche, laut eigener Aussagen, systemische Ansätze im Rahmen von Evaluation bereits angewendet haben, schließt jedoch auch viele positive Reaktionen ein.

5 CHANCEN UND GRENZEN SYSTEMISCHER EVALUATIONSANSÄTZE

Mit Hilfe der ExpertInnen-Interviews erhärtet sich die Erkenntnis, dass die InterviewpartnerInnen unterschiedliche Verständnisse von systemischen Evaluationsansätzen besitzen. Entlang dieser differenten Sichtweisen werden auch die Chancen und Grenzen systemischer Evaluationsansätze unterschiedlich gesehen. So wird von einigen Gesprächspersonen der Standpunkt vertreten, dass die Frage nach den Chancen und Grenzen zu früh gestellt wird, da systemische Verfahrensweisen in der Evaluation bislang allenfalls punktuell angewendet wurden:

„Systemische Ansätze ganz explizit anzuwenden, ist noch nicht der Standard in der Entwicklungszusammenarbeit. Insofern kann man eigentlich nicht schon von Erfahrungsschatz sprechen oder sagen, da kann man Schlussfolgerungen ziehen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Um insbesondere die Chancen abschätzen zu können, bedarf es daher noch einiger Erfahrung und Zeit:

„Man kann nicht sagen, hier haben wir fünf Evaluierungen und das war der Mehrwert des systemischen Denkens“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Chancen	Grenzen
Initiierung von Lern- und Reflexionsprozessen	schwierige Vorabschätzung der Kosten und des Mehrwerts
situationsangepasste Erfassung von Komplexität	divergierende Weltbilder und unzureichende Kooperationsbereitschaft
Steigerung der Kooperationsbereitschaft der StakeholderInnen	geringe Pointiertheit der Evaluationsergebnisse
Steigerung der Akzeptanzfähigkeit und Glaubwürdigkeit von Entscheidungen	Lerninteresse versus Legitimation

Tab. 2: Chancen und Grenzen systemischer Evaluationsansätze; *Quelle: eigene Darstellung*

Daraus kann abgeleitet werden, dass die InterviewpartnerInnen größtenteils auf Erfahrungen aus der systemischen Organisationsberatungspraxis oder anderen Bereichen zurückgreifen, um systemische Evaluationsansätze überhaupt einschätzen zu können. Die in diesem Zusammenhang diskutierten Chancen und Grenzen werden in der folgenden Tabelle nun in kurzer Form veranschaulicht und in den anschließenden Kapiteln genauer dargelegt.

5.1 Chancen: Was man sich von systemischen Evaluationsansätzen verspricht

Grundlegend wird von allen InterviewpartnerInnen die Meinung vertreten, dass sich hinter systemischen Evaluationsansätzen ein großes Potenzial verbirgt. Die Pro-Argumentation bringt dabei nicht nur zum Ausdruck, mit welcher Brille auf systemische Evaluationsansätze geschaut wird, sondern auch, was gegenwärtige Evaluationsdesigns anscheinend nicht zu leisten vermögen. So gesehen stellen die mit systemischen Ansätzen in Verbindung gebrachten Chancen in gewisser Hinsicht eine Kritik am aktuellen Evaluationssystem dar.

INITIIERUNG VON LERN- UND REFLEXIONSPROZESSEN

Die InterviewpartnerInnen verbinden mit systemischen Evaluationsansätzen insbesondere die Chance Lern- und Reflexionsprozesse für die an einer Evaluation Beteiligten zu initiieren:

„Für uns wäre das die Verbesserung des einzelnen Projekts oder Programms einerseits, aber dann auch auf einer höheren Ebene die Verbesserung unserer organisationalen Praxis. Also organisationales Lernen auf Grundlage von

Evaluierungsergebnissen. Evaluierung soll eine Art von Wissen generieren oder zusammenfassen und dann Lernen auf verschiedenen Ebenen, also im Projekt selbst, im Sektor selbst oder für die Organisation ermöglichen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Für eine Person stellt die Initiierung von Lern- und Reflexionsprozessen hingegen kein Spezifikum für systemische Ansätze dar, da auch andere Vorgehensweisen, wie z. B. die der Hermeneutik, Lernen befördern können. Alle weiteren Gesprächspersonen sehen in dem gesteigerten Lerninteresse der AkteurInnen in der Entwicklungszusammenarbeit einen ausschlaggebenden Faktor, weshalb systemischen Ansätzen aktuell eine größere Aufmerksamkeit zuteilwird und auch eine etablierte Anwendung in der Zukunft durchaus zu erwarten ist. In diesem Zusammenhang finden auch geläufige Evaluationsansätze Erwähnung, die genau das in der Vergangenheit nicht leisten konnten:

„Ich glaub der Höhepunkt von dem Ansatz der rigorous impact evaluation ist vorbei. Die sind immer noch groß, aber die sind schon wieder im Abschwung, weil man ja natürlich auch die Schwächen gesehen hat. Bei uns ‚in‘ ist es z. B. so, wir fokussieren im Moment ganz stark auf die Lernfunktion von Evaluierung und weniger auf Rechenschaftslegung. Das heißt für uns sind solche systemischen Ansätze hoch relevant“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Eng mit dem Lerninteresse verknüpft steht die Chance, Dinge kritischer zu hinterfragen und bisher gefahrene Wege in Frage zu stellen. Auf diese Weise können Strategien überdacht oder revidiert werden und es wird Klarheit darüber geschaffen, wie strategisch fortgefahren werden kann.

SITUATIONSANGEPASSTE ERFASSUNG VON KOMPLEXITÄT

Weiterhin assoziieren alle InterviewpartnerInnen mit systemischen Evaluationsansätzen den Vorteil einer der jeweiligen Situation angemessenen Erfassung und Berücksichtigung komplexer Zusammenhänge. Eine Gesprächsperson konsolidiert dies sogar als das Spezifische und als den Mehrwert systemischen Arbeitens schlechthin und stellt in Kontrast zu anderen InterviewpartnerInnen das Lernpotenzial in den Hintergrund. Demzufolge stechen systemische Evaluationsansätze gegenüber anderen Vorgehensweisen durch ihre Leistung hervor, das Gesamtverständnis zu erhöhen bzw. eine gesteigerte Reichhaltigkeit an Informationen zu erzielen. Der Blickwinkel, der über systemische Herangehensweisen ins Auge gefasst wird, ist weiter und schließt auch eher das Unerwartete mit ein. Gerade das Unerwartete, Unplanbare ist das, was den Alltag in der Entwicklungszusammenarbeit auszeich-

net und was in Evaluationsdesigns, gemäß aller InterviewpartnerInnen, viel mehr Berücksichtigung erlangen muss:

„Wenn man mit einer falschen Fragestellung rangeht und die stur durchzieht, also bei der EU in Brüssel ist das eher so die Tendenz, dann kann sein, dass man eine Reihe von Sachen gar nicht wahrnimmt, die auch relevant sind. Und das Systemische ist ja diese Flexibilität und diese Offenheit. Man hat viel mehr Informationen und kann die auch viel angepasster erfassen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Dennoch stellen auch die mittels systemischer Ansätze offengelegten Sichtweisen weiterhin Komplexitätsreduktionen dar, wie eine Gesprächsperson betont. So gesehen wird lediglich eine andere Form der Vereinfachung zur Verfügung gestellt, wobei die vorgenommenen Grenzziehungen sowie die damit einhergehenden Konsequenzen gezielt zur Diskussion gestellt und damit ins Bewusstsein geholt werden. Das bedeutet, dass die jeweiligen Grenzziehungen die Qualität und Gültigkeit der generierten Ergebnisse bedingen und sich in den Ansprüchen, die an die Evaluation herangetragen werden, widerspiegeln müssen:

„Es gibt keine ganzheitliche Herangehensweise. Grenzen werden immer gezogen. Man sollte sich eben der Tragweite dieser Grenzziehung und auch der dahinterliegenden und damit verbundenen Motive und Interessen bewusst sein, klarer herausarbeiten und auch benennen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

STEIGERUNG DER KOOPERATIONSBEREITSCHAFT

Des Weiteren sehen systemische Evaluationsansätze eine wesentlich stärkere Einbeziehung der an einer Evaluation Beteiligten vor. Dieser Weg ist nicht nur konstitutiv für eine gesteigerte Reichhaltigkeit gewonnener Informationen, sondern auch für die Bewältigung zahlreicher Hürden, welche sich bei einer konventionellen Arbeitsweise für EvaluatorenInnen stellen:

„Vor langer Zeit sprach man immer von Evaluationsstarre. Wenn ein Evaluationsteam in ein Projekt kommt, dann zucken alle zusammen und haben Angst: ‚Ah, jetzt kommt das Prüfungsteam, was von der Zentrale geschickt worden ist‘. Wenn man systemisch vorgeht, bezieht man die Leute ein. Dann versucht man auch alle Perspektiven, alle Interessenkonstellationen einzubeziehen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Mittels einer dialogfokussierten Gestaltung von Evaluation wird nicht nur relevanten AkteurInnen die Möglichkeit gegeben, sich mit ihrem Wissensrepertoire und Erfahrungswerten einzubringen, sondern es findet gleichzeitig eine Wertschätzung der AkteurInnen statt, die mit einer erhöhten Bereitschaft, Informationen für EvaluatorInnen Preis zu geben, einhergeht. Hierfür ist ein sensibles Gespür seitens der EvaluatorInnen gefragt, unterschiedliche und eventuell konfligierende Betrachtungsweisen und Interessen gleichwertig anzuerkennen und existierende Konfliktpunkte für alle AkteurInnen sichtbar und damit eventuell lösbar zu machen. Durch die dialogbasierte Berücksichtigung und Offenlegung komplexer zirkulärer Zusammenhänge und der daraus resultierenden Bewusstwerdung der AkteurInnen, dass sie selber Teil des Evaluationssystems sind, wird sowohl die Kooperationsbereitschaft als auch die lösungsorientierte Kommunikation zwischen den Beteiligten gefördert.

STEIGERUNG DER AKZEPTANZFÄHIGKEIT UND GLAUBWÜRDIGKEIT VON ENTSCHEIDUNGEN

Im Kontext einer systemisch angelegten Evaluation werden Entscheidungen möglichst als Resultat eines offenen und intensiven Dialogprozesses getroffen. Damit wird letztlich auch ein ausschließlich vom Evaluationsteam konzipierter Evaluierungsbericht samt Handlungsempfehlungen für die evaluierten AkteurInnen obsolet. Vielmehr macht es Sinn, sich „alle stakeholder mit ins Boot zu holen“ (Interview mit einem Evaluator/einer Evaluatorin) und die dort bestehenden Ressourcen zur eigenen Lösungsfindung zu suchen und zu wecken:

„Ich sage immer, das, was ich als Pflichtübung sehe, sind bei Evaluationen immer Empfehlungen am Schluss. Ohne das ist eine Evaluation nicht vollständig. Die Empfehlungen beziehen sich auf Aktivitäten, die die Leute in dem Projekt ja später umzusetzen haben oder umsetzen sollen. So gesehen macht es Sinn, diese gemeinsam mit den Leuten zu erarbeiten und nicht so im stillen Kämmerlein als Evaluationsteam sich was auszudenken und dann zu sagen, so müsst ihr das jetzt machen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Wie das Zitat implizit untermauert, steigt mit der Einbindung relevanter AkteurInnen auch die Akzeptanz von getroffenen Entscheidungen und künftigen Veränderungen. Es wird nachvollziehbar, vor welchem Hintergrund Entscheidungen so und nicht anders gefällt werden und durch eine gezielte Mitbestimmung werden die evaluierten AkteurInnen angeregt, z. B. Handlungsempfehlungen auch anzunehmen und später in die Tat umzusetzen. Auf diesem Weg fußen getroffene Entscheidungen nicht nur auf einem größeren Fundament, sondern gewinnen gleichzeitig an Authentizität und brechen die häufig beobachteten Muster einer ‚Evaluationsstarre‘ auf.

5.2 Grenzen: Was die Anwendung systemischer Evaluationsansätze scheinbar erschwert

Wie die InterviewpartnerInnen darlegen, sind systemischen Evaluationsansätzen eine Reihe an (bisher noch oft unbeachteten) Chancen immanent, doch gelangen ebenso die Schwächen ins Blickfeld, die nun in den nachstehenden Abschnitten genauer behandelt werden.

SCHWIERIGE VORABEINSCHÄTZUNG DES MEHRWERTS UND DER KOSTEN

Während einerseits von den InterviewpartnerInnen positiv herausgestellt wird, dass systemische Evaluationsansätze besonders dem Prozesscharakter und damit der Komplexität von Projekten und Programmen gerecht werden, wird andererseits die damit einhergehende schwierige Vorausplanung von Nutzen und Kosten problematisiert. Die Anwendung systemischer Ansätze in der Evaluation setzt damit eine hohe Flexibilität der AuftraggeberInnen voraus, die in der Praxis bis dato nur selten anzutreffen ist:

*„Die Schwierigkeit auf institutioneller Seite, wesentlich der Auftraggeberseite, liegt darin, einerseits einen Vertrag mit dem Evaluationsteam abzuschließen, wo die Aufgabenstellung festgelegt werden muss und auf der anderen Seite prozessabhängig möglichst hohe Flexibilität beizubehalten, um dann Anpassungsprozesse vornehmen zu können. Da sehe ich die Schwierigkeit auf der institutionellen Ebene, wenn man so einen Vertrag festklopfen muss“
(INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).*

Diese Problematik wird in allen Interviews als große Hürde herausgestellt, die zuerst überwunden werden muss, möchten systemische Evaluationsansätze ernsthaft implementiert werden. Es wird der Standpunkt vertreten, dass das Verständnis für mehr Flexibilität auf der AuftraggeberInnen-Seite erst noch wachsen muss und es bis zu dieser Einsicht noch etwas Zeit benötigt. Einige InterviewpartnerInnen erkennen bereits einen sehr langsam voranschreitenden Wandel in den Denkweisen der AuftraggeberInnen, die zu mehr Offenheit und Flexibilität tendieren und damit systemischen Evaluationsansätzen den Weg zur Erprobung ebnen. Trotz zuversichtlicher Aussichten steht diese Entwicklung noch ganz am Anfang.

Im Großen und Ganzen wird die Anwendung systemischer Evaluationsansätze in der Praxis jedoch nur realisierbar sein, wenn insbesondere die AuftraggeberInnen ein Interesse am Überraschenden mitbringen sowie eine Akzeptanz gegenüber

einer sehr flexiblen Herangehensweise. Sind diese Voraussetzungen nicht gegeben, ist eine gewinnbringende Umsetzung schlichtweg unmöglich.

DIVERGIERENDE WELTBILDER UND UNZUREICHENDE KOOPERATIONSBEREITSCHAFT

Letztendlich kann die Qualität der Evaluationsergebnisse nur so gut sein, wie es das Zusammenspiel zwischen EvaluatorInnen und Evaluierten inklusive AuftraggeberInnen hergibt. Neben der Flexibilität auf Seiten der AuftraggeberInnen müssen daher auch alle anderen an dem Evaluationsvorhaben Beteiligten bestimmte Voraussetzungen mitbringen. Hierzu zählt unter anderem ein Weltverständnis, nachdem es unterschiedliche Meinungen und Perspektiven zu einem Sachverhalt geben kann, die alle gleichberechtigt und gleichwertig nebeneinanderstehen. Dies erfordert eine hohe Bereitschaft, seine eigenen Ansichten und Interessen zu kommunizieren und auch denen der anderen ebenbürtig Gehör zu schenken. Allen Beteiligten wird ein hohes Maß an Eigeninitiative und Selbstreflexivität abverlangt, wenn die Entscheidungsfindung auf einer breiten Basis fußen soll. Ist diese Einstellung nicht vorhanden, ist die gewinnbringende Umsetzung gefährdet.

Laut einer Gesprächsperson ist die Voraussetzung einer großen Kooperationsbereitschaft und Toleranz insbesondere dann gegeben, wenn ein offenes Lerninteresse besteht. Allerdings liegt genau darin auch ein wesentlicher Knackpunkt, denn nicht jede Organisation möchte lernen. Stattdessen herrscht noch häufig ein Weltbild vor, welches die Vielfältigkeit von Sichtweisen und die Komplexität von Zusammenhängen ausklammert und somit mit der Anwendung systemischer Evaluationsansätze konterkariert. An dieser Stelle stoßen daher systemische Evaluationsansätze eindeutig an Grenzen:

„Ich habe Kollegen, die kommen aus der Wissenschaft und sind davon überzeugt, dass die Welt objektiv messbar ist. Das ist auch nichts, worüber man streiten kann. Das sind philosophische Grundannahmen oder Glaubenssätze. Das ist dann halt so“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Daraus ableitend kann konstatiert werden, dass die Entscheidung für die Verwendung systemischer Evaluationsansätze immer auch mit Blick auf die beteiligten AkteurInnen erfolgen muss, um einen Nutzen und vor allem auch eine Glaubwürdigkeit der Ergebnisse zu gewährleisten. Stößt ein systemisches Weltbild bei den Beteiligten auf größeren Widerstand, empfiehlt es sich logischerweise nicht, systemische Ansätze anzuwenden oder zumindest, wie von den InterviewpartnerInnen kommentiert, offensiv mit systemischem Vokabular zu operieren. Für den Bereich der Entwicklungszusammenarbeit kann jedoch festgehalten werden, dass sich hier viele BefürworterIn-

nen des systemischen Denkens finden, die an dem Bedeutungsgewinn systemischer Evaluationsansätze sehr interessiert sind und die Entwicklung auch stark mittragen.

Die erläuterten Fähigkeiten und Grundeinstellungen für eine systemische Vorgehensweise sind nicht als selbstverständlich anzusehen und stellen auch die EvaluatorInnen vor große Herausforderungen. So verharren viele EvaluatorInnen noch immer in der Rolle von KontrolleurInnen, die beabsichtigen, objektives Recht zu sprechen. Hingegen gilt es „mehrere Hüte im Laufe der Evaluation aufsetzen“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN) zu können und insbesondere durch eine Reflexions- und Moderationsfähigkeit zu überzeugen. Gemäß einer Gesprächsperson seien diese Hürden weniger in den Fähigkeiten der EvaluatorInnen zu suchen:

„Die Grenzen sind da eher im Kopf und nicht in den Fähigkeiten. So erlebe ich das“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

GERINGE POINTIERTHEIT DER EVALUATIONSERGEBNISSE

Einige InterviewpartnerInnen erachten als Nachteil, dass sich die mittels systemischer Ansätze generierten Evaluationsergebnisse häufig schlecht verkaufen lassen aufgrund des immanenten hohen Komplexitätsgrades. Wie bereits an einer anderen Stelle erwähnt, verkehrt sich damit ein wesentlicher Vorteil gleichzeitig zu einem Nachteil:

„Es ist auch komplex. Wenn man systemisch arbeitet, wird, wenn man die ganzen Zwischenverbindungen und Auswirkungen und Einflussfaktoren sieht, die ganze Angelegenheit recht komplex. Man nähert sich vielleicht der Situation eher an, aber es ist natürlich auch nicht so pointiert“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Der hohe Komplexitätsgrad wird dann als problematisch betrachtet, wenn die Beteiligten systemische Grundannahmen nicht teilen und von einer einfach abzubildenden und objektiv messbaren Realität ausgehen:

„Sie können sich zum Teil nicht so gut verkaufen, weil manche Auftraggeber eine Klarheit und Eindeutigkeit wollen. Die Welt ist nun mal nicht so eindeutig und mit systemischen Ansätzen kann man so eine Eindeutigkeit nicht so leicht vorgaukeln“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Alle InterviewpartnerInnen nehmen speziell auf dem Gebiet der Politik große Schwierigkeiten im Umgang mit komplexen Evaluationsergebnissen wahr. Dort liegen eine andere Handlungslogik und ein Weltverständnis vor, welche nach

einer stark reduzierten Wirklichkeit verlangen und alternative Betrachtungsweisen schlichtweg ausblenden:

„Sie haben zum Schluss eben nicht eine Zahl und sie haben auch nicht, es hat funktioniert oder nicht, sondern sie haben eher eine story. Wie bereitet man denn sowas auf für einen Entscheidungsträger, der maximal eine Seite liest? Das sind alles noch so ungelöste Fragen und das ist nicht das, was die Politik unbedingt braucht und will im Moment“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

In diesem Zusammenhang verweist eine Gesprächsperson darauf, dass die Umsetzung systemischer Ansätze als auch die Komplexität der Evaluationsergebnisse häufig falsch eingeschätzt werden. Demzufolge liefern systemische Evaluationsansätze ebenfalls „eine andere Form der Vereinfachung“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN), weshalb die oft vertretene Erwartungshaltung nach einem äußerst komplizierten, teuren und langwierigen Vorgehen den Kern nicht trifft. Stattdessen steht im Mittelpunkt, mit welchen Einschränkungen und Grenzziehungen innerhalb eines Evaluationsvorhabens operiert wird, um darauf aufbauend die Konsequenzen für die Gültigkeit und Qualität der Evaluationsergebnisse abzuleiten.

LERNINTERESSE VERSUS LEGITIMATION

Anhand der ExpertInnen-Interviews kommt zum Vorschein, dass systemische Evaluationsansätze sich besonders gut für Situationen eignen, in denen bei den Beteiligten ein offenes Lerninteresse besteht. Soll Evaluation hingegen der Funktion einer Rechenschaftslegung und Legitimation nachkommen, schließen sich Lerneffekte aus oder zumindest wird ihre Verwirklichung dann erheblich erschwert:

„Je mehr Auftraggeber ein offenes Lerninteresse haben, desto offensiver und desto besser kann man auch systemische Ansätze einbringen. Je mehr es politisch hoch gepuschte Legitimationserwartungen zu befriedigen gilt, desto stärker muss man sich natürlich auch auf solche Aspekte konzentrieren“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Auf die Frage, warum die beiden Evaluationsfunktionen der Legitimation und des Lernens eher in einem konfligierenden Verhältnis zueinander stehen, rezipiert dieselbe Gesprächsperson Folgendes:

„Die wollen sozusagen harte Fakten, die sie dann in Broschüren tun. Das sind Fakten, die vielleicht so wie Fakten aussehen, aber aus systemischer Sicht nicht wirklich Fakten sein können. Also solche Evaluierungen mache ich selber

eigentlich nicht. Man kennt die Szene ja und der Druck ist schon relativ hoch, gerade auch im internationalen Feld genau solche Dinge abzuliefern und da passt das systemische Weltbild nicht wirklich rein“ (INTERVIEW MIT EINEM EVALUATOR/EINER EVALUATORIN).

Den beiden voranstehenden Zitaten zufolge stoßen systemische Evaluationsansätze an Grenzen, wenn mit einer Evaluation neben einem Lerninteresse ebenso eine hohe Legitimationserwartung verbunden wird. Diese Erkenntnis steht in Einklang mit der bereits beleuchteten Ansicht, dass auf politischer Ebene andere Handlungslogiken und Weltanschauungen vorherrschen, die eine starke Vereinfachung komplexer Zusammenhänge einfordern und entsprechend systemische Evaluationsansätze wenig affirmieren. Wird sich vor Augen geführt, dass innerhalb der Entwicklungspolitik vor allem das Ziel verfolgt wird, Aktivitäten der Entwicklungszusammenarbeit nach außen hin zu legitimieren, so wird einmal mehr nachvollziehbar, warum das Lernen als Evaluationsfunktion einen nachrangigen Stellenwert einnimmt und bislang ein relativ geringes Interesse an systemischen Evaluationsansätzen zu beobachten war.

Trotz des Spannungsverhältnisses zwischen Lernen und Rechenschaftslegung im Rahmen von Evaluation gehen einige Gesprächspersonen davon aus, dass sich in der deutschen Entwicklungszusammenarbeit systemische Evaluationsansätze künftig stärker etablieren werden. Zum einen wird den deutschen Durchführungsorganisationen bei der Wahl der Evaluationsmethoden ein relativ großer Spielraum eingeräumt und zum anderen begünstigt die Gründung des DEval eine stärkere Fixierung der Durchführungsorganisationen auf die Lernfunktion von Evaluation. Solange der Beitrag zur Legitimation geleistet wird, besteht kein Anlass, die Anwendung anderweitiger Evaluationsansätze, auch entgegen der Priorisierungen entwicklungspolitischer AkteurInnen, zu blockieren.

6 ZUSAMMENFASSUNG UND FAZIT

Das Ziel dieses Beitrages bestand darin, Kommunikationsprozesse über systemische Evaluationsansätze in der deutschen staatlichen Entwicklungszusammenarbeit zu beleuchten und darüber hinaus die mit systemischen Evaluationsansätzen verknüpften Chancen und Grenzen aufzuzeigen. Dabei offenbarte die Auswertung von fünf ExpertInnen-Interviews sowohl einen allgemeinen Konsens der InterviewpartnerInnen im Verständnis über systemische Evaluationsansätze, als auch große Ambivalenzen und Unterschiede.

Alle InterviewpartnerInnen brachten dabei zur Sprache, dass eine allgemeine Definition zu systemischen Evaluationsansätzen bis dato fehlt und sich daher die Charakterisierung und Abgrenzung äußerst schwierig gestalten. Die diesbezüglichen Unsicherheiten haben bei den meisten Gesprächspersonen die Ansicht reifen lassen, dass sich in der Evaluationsforschung bislang noch nicht ausreichend dem Thema gewidmet wurde. Vor dem Hintergrund, dass die InterviewpartnerInnen ganz unterschiedliche Verständnisse von systemischen Evaluationsansätzen besitzen, wurden auch die Chancen und Grenzen unterschiedlich diskutiert. Jedoch herrschte ein allgemeiner Konsens darüber, dass zum jetzigen Zeitpunkt von keinem verlässlichen Erfahrungsschatz ausgegangen werden, geschweige denn eine Bilanz zu systemischen Evaluationsansätzen gezogen werden kann. Die von den ExpertInnen exemplifizierten Chancen und Grenzen beruhen letztlich auf dem Erfahrungsschatz aus der systemischen Organisationsberatung und stellen folglich eher Erwartungen an systemische Evaluationsansätze dar.

Mit systemischen Evaluationsansätzen wurde vor allem die Chance der Initiierung von Lern- und Reflexionsprozessen assoziiert, von denen alle an einem Evaluationsvorhaben Beteiligte profitieren können. Hierunter wurde organisationales Lernen subsumiert sowie auch die Verbesserung einzelner Projekte und Programme. Des Weiteren werden mittels systemischer Ansätze in der Evaluation komplexe Zusammenhänge situationsangepasster und damit angemessener erfasst. Das Gesamtverständnis wird erhöht, da viel mehr Faktoren jenseits von Evaluationsfragestellungen und Indikatoren miteinbezogen werden, die von Relevanz sind und bei anderen Ansätzen unberücksichtigt bleiben. Der Blickwinkel, der über systemische Herangehensweisen ins Auge gefasst wird, schließt eher das Unerwartete mit ein, was dem Charakter der Entwicklungszusammenarbeit und dem Wunsch nach prozessorientierter Bewertung gerecht wird. Eine weitere Chance besteht darin, dass sich durch die stärkere Einbeziehung der an einer Evaluation Beteiligten die Kooperationsbereitschaft erhöht und damit einer ‚Evaluations-

starre‘ entgegengewirkt wird. Dadurch nimmt nicht nur die Reichhaltigkeit der Informationsgrundlage zu, sondern auch die Akzeptanzfähigkeit von Evaluationsergebnissen und daraus abgeleiteten Entscheidungen.

Die InterviewpartnerInnen sahen die Anwendung systemischer Evaluationsansätze allerdings auch mit einigen Grenzen verknüpft. Demnach lässt sich im Voraus einer systemisch angelegten Evaluation nicht sagen, welche genauen Kosten und vor allem welcher Mehrwert respektive Nutzen sich aus diesen Vorgehensweisen ergibt. Dies stellt speziell AuftraggeberInnen vor neue Hürden. Gehen diese mit einem derart flexiblen und offenen Vorgehen nicht konform und stellen zudem sehr hohe Erwartungen an die Erzielung eines möglichst raschen Nutzens, ist die gewinnbringende Anwendung und Etablierung systemischer Evaluationsansätze stark gefährdet. Dabei wird nicht nur von AuftragerInnen, sondern auch von allen Beteiligten, einschließlich den EvaluatorInnen, ein hohes Maß an Eigeninitiative, Selbstreflexivität und Offenheit gegenüber anderen Denkweisen und Perspektiven abverlangt, was in der Praxis nicht immer als selbstverständlich angesehen werden kann. Die Qualität der Evaluationsergebnisse und damit der Nutzen einer systemisch angelegten Evaluation können damit letztendlich nur so gut sein, wie es das Zusammenspiel zwischen AuftraggeberInnen, EvaluatorInnen und Evaluierten hergibt. Als weitere Hürden wurden der hohe Komplexitätsgrad und damit die geringe Pointiertheit der Evaluationsergebnisse angeführt. Viele AuftraggeberInnen scheuen daher vor systemischen Evaluationsansätzen zurück, da sie klare und eindeutige Ergebnisse erwarten. Steht in diesem Zusammenhang statt eines Lerninteresses das Ziel der Legitimation im Mittelpunkt, können systemische Evaluationsansätze tatsächlich an Grenzen stoßen, da sie eine andere Weltanschauung und Handlungslogik bedienen, als auf politischer Ebene vorherrschen.

Zusammenfassend kann aus den beleuchteten Chancen und Grenzen abgeleitet werden, dass die Entscheidung für die Verwendung systemischer Evaluationsansätze immer auch mit Blick auf die beteiligten AkteurInnen und den gegebenen Rahmenbedingungen erfolgen muss, um einen Nutzen und vor allem auch eine Glaubwürdigkeit der Ergebnisse sicherzustellen. Ein offenes Lerninteresse bspw. ist für eine gewinnbringende Umsetzung systemischer Evaluationsansätze unerlässlich sowie auch eine Offenheit gegenüber ‚neuen‘ bzw. unkonventionellen Herangehensweisen und Methoden im Rahmen von Evaluationsvorhaben. Der Mehrwert systemischer Evaluationsansätze kann derzeit nicht verlässlich eingeschätzt werden, da die ersten Erprobungen in der Praxis bedingt durch ein sehr dünnes wissenschaftliches Fundament mit großen Unsicherheiten und Hürden einhergehen. Hier kann die Anwendung systemischer Evaluationsansätze an Grenzen stoßen, wenn die Erwartungen an einen möglichst raschen Nutzen dominieren, während das Lerninteresse im Hintergrund verbleibt. Angesichts dieser Referenzfolie liegt es vor allem auch in den Händen der Wissenschaft, systemische Evaluationsansätze zu erforschen, authentisch zu vertreten und damit in das Sichtfeld von potenziellen AuftragerInnen von Evaluationsvorhaben zu rücken.

Ohne einen wesentlichen Rückhalt aus den Reihen der Wissenschaft erscheint daher eine bewusste und verbreitete Anwendung systemischer Evaluationsansätze in der Praxis als äußerst unwahrscheinlich.

LITERATURVERZEICHNIS

- ARBEITSKREIS „EVALUATION VON ENTWICKLUNGSPOLITIK“ IN DER DEGEVAL – GESELLSCHAFT FÜR EVALUATION E. V., AG WIRKUNGSANALYSE (2009): Verfahren der Wirkungsanalyse. Eine Landkarte für die entwicklungspolitische Praxis (abrufbar unter URL: http://www.entwicklung.at/uploads/media/Verfahren_der_Wirkungsanalyse.pdf, Stand: 02.02.2015).
- BMZ – BUNDESMINISTERIUM FÜR WIRTSCHAFTLICHE ZUSAMMENARBEIT UND ENTWICKLUNG (2005): Wirksamkeit der Zusammenarbeit. Die Erklärung von Paris. („Paris Declaration on Aid Effectiveness“) (abrufbar unter URL: http://www.bmz.de/de/was_wir_machen/ziele/ziele/parisagenda/paris/index.html, Stand: 02.02.2015)
- BMZ – BUNDESMINISTERIUM FÜR WIRTSCHAFTLICHE ZUSAMMENARBEIT UND ENTWICKLUNG (2010) (Hrsg.): Die Millenniums-Entwicklungsziele. Hintergründe – Zielerreichung – Engagement (abrufbar unter URL: http://www.bmz.de/de/mediathek/publikationen/reihen/infobroschueren_flyer/infobroschueren/Materialie204_Informationsbroschuere_04_2010.pdf, Stand: 02.02.2015).
- BORRMANN, A. (2009): Entwicklungszusammenarbeit. Evaluation in der Entwicklungszusammenarbeit vor neuen Herausforderungen (abrufbar unter URL: <http://www.hwwi.org/publikationen/update/hwwi-update-einzelansicht/evaluation-in-der-entwicklungszusammenarbeit-vor-neuen-herausforderungen/6709.html>, Stand: 02.02.2015).
- BORRMANN, A.; STOCKMANN, R. (2009a): Evaluation in der deutschen Entwicklungszusammenarbeit. Band 1. Systemanalyse. Münster: Waxmann.
- BORRMANN, A.; STOCKMANN, R. (2009b): Evaluation in der deutschen Entwicklungszusammenarbeit. Band 2. Fallstudien. Münster: Waxmann.
- DEVAL – DEUTSCHES EVALUIERUNGSINSTITUT FÜR ENTWICKLUNGSZUSAMMENARBEIT (2012) (Hg.): Kurzmeldung vom 19.12.2012: DEVal-Expertendiskussion „Dealing with Complexity“ mit Bob Williams und Richard Hummelbrunner am 04.12.2012 (abrufbar unter URL: <http://www.deval.org/de/kurzmeldung-detail/deval-expertendiskussion.html>, Stand: 02.02.2015).

- FAUST, J. (2010): Entwicklungspolitik. Wirkungsevaluierung in der Entwicklungszusammenarbeit. In: bpb – Bundeszentrale für politische Bildung (Hrsg.): Aus Politik und Zeitgeschichte, Ausgabe 10/2010 (abrufbar unter URL: <http://www.bpb.de/apuz/32914/wirkungsevaluierung-in-der-entwicklungszusammenarbeit?p=all>, Stand: 02.02.2015)
- FROSCHAUER, U. (2003): Wollen Sie Legitimation oder Lernen? Prof. Dr. Ulrike Froschauer vom Institut für Soziologie der Universität Wien über unterschiedliche Evaluierungsansätze und Erfolgsvoraussetzungen (abrufbar unter URL: <http://www.leaders-circle.at/evaluierung1.html>, Stand: 02.02.2015).
- GTZ – DEUTSCHE GESELLSCHAFT FÜR TECHNISCHE ZUSAMMENARBEIT GMBH, STABSSTELLE UNTERNEHMENSENTWICKLUNG, GRUPPE POLITIK UND STRATEGIE (2010) (Hg.): Scaling-up – Breitenwirksame Entwicklungszusammenarbeit. Eine Orientierung für die Praxis (abrufbar unter URL: http://www.giz.de/Themen/de/dokumente/Orientierung_Scaling-up-de.pdf, Stand: 15.07.2013).
- HUMMELBRUNNER, R. (2011): Systems thinking and evaluation. In: Evaluation. 17. Heft 4, S. 395–403 (abrufbar unter URL: <http://evi.sagepub.com/content/17/4/395.extract>, Stand: 02.02.2015).
- KEVENHÖRSTER, P.; VAN DEN BOOM; D. (2009): Entwicklungspolitik. Wiesbaden: VS Verlag für Sozialwissenschaften.
- KLINGEBIEL, S. (2013): Entwicklungszusammenarbeit – eine Einführung. Studies 73. Deutsches Institut für Entwicklungspolitik (abrufbar unter URL: http://www.die-gdi.de/uploads/media/Studies_73.pdf, Stand: 02.02.2015).
- KRANZ-PLOTE, J. (2006): Der Millenniumsprozess und die deutsche Entwicklungspolitik. BMZ Artikel für die AGEV-Veranstaltung „Die Millenniums-Entwicklungsziele: Entstehung, Kritik und Perspektiven“ im März 2006 (abrufbar unter URL: http://www.ageg.de/_ageg/news-events/MDG-Artikel_BMZ.pdf, Stand: 02.02.2015)
- KÜHL, S. (2008): Das Evaluations-Dilemma der Beratung: Evaluation zwischen Ansprüchen von Lernen und Legitimation. In: Postheroisches Management. Heft 3/2008, S. 64–71 (abrufbar unter URL: <http://www.yumpu.com/de/document/view/8825475/das-evaluations-dilemma-der-beratung-31-osterreichischen->, Stand: 02.02.2015).

- LAMNEK, S. (2005): *Qualitative Sozialforschung. Lehrbuch* (4. Aufl.). Weinheim/Basel: Beltz Verlag.
- MEYER, W. (2004): *Indikatorenentwicklung. Eine praxisorientierte Einführung* (2. Aufl.). CEval Arbeitspapiere. Nr. 10 (abrufbar unter URL: http://www.ceval.de/typo3/fileadmin/user_upload/PDFs/workpaper10.pdf, Stand: 02.02.2015).
- NUSCHELER, F. (2004): *Lern- und Arbeitsbuch Entwicklungspolitik* (5. Aufl.). Bonn: Dietz Verlag.
- NUSCHELER, F. (2008): *Die umstrittene Wirksamkeit der Entwicklungszusammenarbeit*. Duisburg: Institut für Entwicklung und Frieden, Universität Duisburg-Essen (INEF-Report 93/2008) (abrufbar unter URL: <http://inef.uni-due.de/page/documents/Report93.pdf>, Stand: 02.08.2013).
- PATTON, M. Q. (2010): *Developmental Evaluation. Applying Complexity Concepts to Enhance Innovation and Use* (Inhaltsangabe abrufbar unter URL: <http://www.guilford.com/books/Developmental-Evaluation/Michael-Quinn-Patton/9781606238721/summary>, Stand: 02.02.2015).
- PEEL, D. (2003): *Town Centre Management: Multi-stakeholder Evaluation. Increasing the Sensitivity of Paradigm Choice*. In: *Planning Theory and Practice*. Band 4, Heft 2, S. 147–164.
- RAUCH, T. (2012): *Entwicklungsstrategien und Entwicklungsinstrument. Strategien und Instrumente nachhaltiger Entwicklung – ein multidimensionaler Mehr Ebenen-Ansatz*. In: Fiege, K.; Rauch, T. (Hg.): *Inhalte und Methoden für ein erfolgreiches Wirken in einem komplexen Berufsumfeld* (1. Aufl.). Berlin: Seminar für Ländliche Entwicklung Publikationsserie, S. 40–82 (abrufbar unter URL: <http://edoc.hu-berlin.de/series/sle/248/PDF/248.pdf>, Stand: 02.02.2015).
- ROLFES, M.; WILHELM, J. L. (2007): *Politische Steuerung und Zielperspektiven einer sozial-integrativen Stadtentwicklung im Bundesland Brandenburg*. In: Bundesamt für Bauwesen und Raumordnung. Akademie für Raumforschung und Landesplanung. *Raumforschung und Raumordnung* Heft 2/2007. Jg. 65: Carl Heymanns Verlag, S. 109–121.
- SILVESTRINI, S. (2011): *Ex-ante-Evaluation. Ein Planungsansatz für die Entwicklungszusammenarbeit*. Sozialwissenschaftliche Evaluationsforschung. Band 10. Münster: Waxmann.

- STOCKMANN, R. (HG.) (2007): Handbuch zur Evaluation. Eine praktische Handlungsanleitung. Sozialwissenschaftliche Evaluationsforschung. Band 6. Münster: Waxmann.
- STOCKMANN, R.; MENZEL, U.; NUSCHELER, F. (2010): Entwicklungspolitik. Theorien – Probleme – Strategien. München: Oldenbourg.
- UN – UNITED NATIONS (HG.) (2005): Offizielle Liste der Indikatoren für die Millenniums-Entwicklungsziele (abrufbar unter URL: <http://www.un.org/depts/german/millennium/MDG-Indikatoren.pdf>, Stand: 02.02.2015).
- WILHELM, J. L. (2012): Wozu Evaluation? Organisationssysteme bewerten Stadtteilmförderung mit Kalkül. Potsdam: Universitätsverlag.
- WILLIAMS, B.; IMAM, I. (HG.) (2006): Systems concepts in evaluation. An expert anthology. American Evaluation Association/Edge Press: Point Reyes (abrufbar unter URL: <http://preval.org/files/Kellogg%20enfoque%20sistematico%20en%20evaluacion.pdf>, Stand: 02.02.2015).

LOST IN COMPLEXITY – HOW TO EVALUATE SUSTAINABLE DEVELOPMENT?

Hendrik Smid

1 BACKGROUND: SUSTAINABLE DEVELOPMENT AND EVALUATION

“The great unexplored frontier is evaluation under conditions of complexity.”
PATTON (2011: 1)

1.1 Sustainable development: controversies, delusions and complexity

Sustainable development is one of the most contested theoretical concepts and real world challenges in recent times. While the original definition of the Brundtland Commission of “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED 1987: 43) is widely accepted, the actual conceptualization, implementation and evaluation of sustainable development is far from any acceptance. There is widespread disappointment about the discourse on sustainable development, which results in delusions, controversies and considerations of sustainable development as ‘oxymoron’ and/or most ‘wicked’ problem of modernity (ROBINSON 2004; VOSS, BAUKNECHT, KEMP 2006; VOSS, KEMP 2005). Since the Brundtland definition, various alternative definitions and interpretations of sustainable development have emerged. While many of these are based upon the so called ‘three-pillar’ or ‘triple bottom line’ concept (POPE, ANNANDALE, MORRISON-SAUNDERS 2004) there is still “no clear agreement as to what sustainable development is” (BOND, MORRISON-SAUNDERS 2011: 2) as well as “what sustainable development should be and how it can be attained by (global) policy” (STÖRMER, SCHUBERT 2007: 2). As to ROTMANS (2006) “sustainable development is an essentially contested notion, because it is inherently complex, normative, subjective and ambiguous” (ROTMANS 2006: 37). Different actors have different understandings of sustainable development, its related problems and how best it can be implemented in practice (BOND, MORRISON-SAUNDERS 2011). Sustainability hence needs to be answered towards the questions of: sustainability for what, for whom, in what way, for how long, where and at what scale? (MONEVA et al. 2006: 124).

As a result, the challenges in reaching sustainable development can be considered as an ‘unstructured problem’. The term was introduced by HOPPE (2011: 73) to characterize problem states which are far from agreement on norms and values at stake and furthermore are far from certainty on required and available knowledge. In unison, ROTMANS describes current societal problems related to sustainable development as ‘persistent problems’ which are “complex, ill-structured, involve many stakeholders, are surrounded by structural uncertainties, and are hard to manage” (ROTMANS 2006: 36). Precisely these features characterize the state of the art in the field of German development cooperation. HUMMELBRUNNER (2011) succinctly claims that sustainable development interventions of German development cooperation are faced by conditions in which:

- » *“their success depends on the interaction of various resources (e. g. physical, economic, social, cultural) – and on the quality of collaboration between key actors responsible for these resources;*
- » *The main actors (funders, implementers, partners, beneficiaries) are social beings with specific needs, interests and values;*
- » *The context in which these interventions take place is often insecure and turbulent, which makes it hard to foresee routes or predict effects;*
- » *Development not only requires transfers (of resources, technology or knowledge) but also is conceived as an open change process, which cannot be determined in advance and needs to be continuously shaped, in order to cope with changes or integrate lessons learned along the way.*

As a result, interventions become multi-layered and multi-faceted, i. e. more complex [which] is the result of many different elements (e. g. actors, actions, factors) and their linkages” (HUMMELBRUNNER 2011: 395).

The key challenge with regard to evaluation is hence to find appropriate ways and frameworks, which are sensitive to such complex, unstructured and persistent problems. In German development cooperation, there is, however, a situation in which “current evaluation approaches are often based on a mechanistic worldview and a linear results chain. But in order to understand which interventions are successful under which conditions, evaluations need to give consideration to the multiple facets of social change” (GIZ 2013: 48). With regard to sustainable development, it is difficult to determine not only the intended effects of sustainable development interventions “but also a range of adverse side-effects” (VOSS, KEMP 2005: 9). According to HUMMELBRUNNER such uncertainties in evaluations can further generate disagreement: “unbalanced or simplistic attempts to reduce this complex picture will not only harm the credibility of an evaluation, but also

bring forth resistance from those who do not feel themselves properly represented” (HUMMELBRUNNER 2007: 161–162). Similarly VOSS and KEMP argue:

“It is not possible, for pragmatic reasons, to rely on simpler models of the causes which are behind sustainability problems. If complexity becomes externalised from cognitive models, the world will still stay as complex as it is and ignored linkages will still be effective. Inadequate problem constructions thus come back in form of unexpected consequences when strategies are implemented in the real world” (VOSS, KEMP 2005: 11).

According to ROTMANS (2006) scientists so far failed to develop adequate tools to comprehensively assess sustainable development in its complexity. He claims:

“Here we run into the ‘sustainability paradox’: the unsustainability problems humankind is faced cannot be solved with current tools and methods that were applied—or seemed to work—in the past. [...] What we can do, however, is using current knowledge, tools and methods in the best possible manner, while developing a new paradigm that better reflects the complexity of sustainable development [...] to support the sustainability decision-making process in a more adequate manner” (ROTMANS 2006: 37).

To sensitively deal with complexity, the integration of ‘systems thinking’ into evaluations can play an important role. Evaluations provide an important contribution for ongoing adaptation, if they are designed for informing strategic management and building reflective capacity of stakeholders (HUMMELBRUNNER 2007). However, in states of complexity, alternative forms of evaluation which are more open and emphasize experimentation and learning are required (VOSS et al. 2006: 6). In this respect, VOSS and KEMP hold that strategies towards adaptive management “should feature experimentation, monitoring and evaluation in order to systematically work with new experiences, altered interpretations and changed circumstances” (VOSS, KEMP 2005: 11–12). Referring to the discourse of evaluators during a joint American Evaluation Association and Canadian Evaluation Society conference in 2005, WILLIAMS and IMAM (2007) state that ‘systems thinking’ and applying systemic approaches in evaluations:

“Makes you think differently; offers more effective ways of dealing with complexity and complex situations; links the local and global, across silos, sectors and disciplines; provides tools to work with different opinions of stakeholders; develops new ways of understanding situations; pays attention to properties that emerge unexpectedly; acknowledges the richness and interdependence of real life” (WILLIAMS, IMAM 2007: 3).

1.2 Towards systemic approaches in evaluation of sustainable development?

Ever since sustainable development has become an overarching goal in international policies and programmes (STÖRMER, SCHUBERT 2007: 1–2) evaluations dealing with sustainable development are rapidly emerging (DE RIDDER et al. 2007). Sustainability evaluations have gained momentum to assist policy makers for more evidence-based decision making and moreover have become a legal prerequisite as of binding European and national law on impact assessments (ARBTER 2007: 85). Evaluations are, hence, crucial processes inhibiting various instruments for knowledge production in the intersection of science and policy, as well as practice. Not only have evaluations been considered as credible evidence base for public policy, accountability, steering and actions towards sustainable development; they, furthermore, have the potential of functioning as a critical element in procedural sustainability with strategic reflection, learning and informed citizen deliberation (BÄCKSTRAND et al. 2010; MEADOWCROFT 2007; MILLER 2012; STOCKMANN 2011; STOKER 1998). On the opposite, however, how to measure sustainable development in its complexity remains largely unclear (HARDI 2007). In response to more accurately capture and deal with the challenge of complexity, it seems the need for systemic approaches in sustainability evaluation has (re-) surfaced.

An online search analysis at the Web of Science for scientific articles with ‘systemic’ and ‘evaluation’ in their title gives an idea of this trend. The findings show that from the period between 1993 and 2012, an overall significant increase of articles using these keywords in their title can be identified. The growth is however slightly wavy with recessions following prior peaks, as can be recognised for the climax in 2010 and a decline for the following years. This is, however, different with regard to citations of articles with ‘systemic’ and ‘evaluation’ in their title. Here, there is an exponential increase with a peak in 2012 (see fig. 1).¹

Raising the agenda on systemic evaluation approaches in German development cooperation was essentially led by the German Evaluation Society (DeGEval, Deutsche Gesellschaft für Evaluation e.V.). The society has established an exclusive task force on systemic evaluation approaches (within the development policy working group) and hosted a DeGEval conference on the same topic in spring 2011. Two years later, a joint discussion paper was launched in yet another DeGEval conference (see HUMMELBRUNNER et al. in this book). Meanwhile, the evaluation and audit division of the Federal Ministry of Economic Cooperation

¹ Note: For 2013 the data just encompassed the first four month and thus shows a significantly lower value. Noteworthy, is further that the publication sources that could be reviewed and identified by the Web of Science databases (though very diverse) are mainly attributed to medical sciences and biology.

and Development (BMZ) and the evaluation unit of GIZ, the German bilateral aid agency, held an international conference on systemic approaches to evaluation in January 2011, which was well attended by numerous participants from development policy, academia, consulting firms and NGOs. In addition, the 16th annual DeGEval evaluation conference dealt with the topic of ‘evaluation and complexity’ in Munich, September 2013.

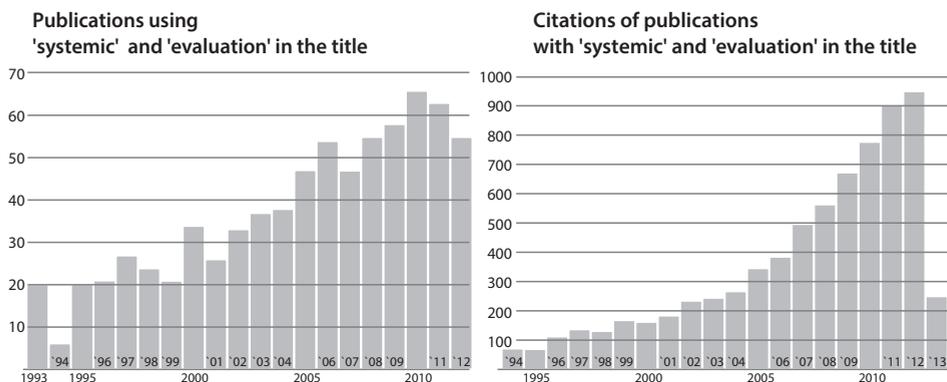


Fig. 1: Web of Science publications and citations of articles using ,systemic' and ,evaluation' in the title
 Source: Web of Science databases (SCI-EXPANDED, SSCI, A&HCI)

Embedded in these dynamics, the present research seeks to enhance the insights on perspectives from various actor groups on systemic evaluation approaches in German development cooperation. By applying an empirical qualitative analysis, this study aims to explore evaluation approaches in sustainable development of German development cooperation and to identify opportunities and limitations for integration of systemic evaluation approaches. The research does not intend to define an ‘ideal’ sustainability evaluation framework, process or policy. Rather, its focus is on filling an emerging research gap at the intersection of sustainability science and evaluation science, that is, sustainability evaluation using systemic approaches in the context of German development cooperation.

2 EMPIRICAL ANALYSIS OF PERSPECTIVES ON EVALUATION IN GERMAN DEVELOPMENT COOPERATION

*“Theory must be tempered with reality.”
(Jawaharlal Nehru, Prime Minister of India
in 1889–1964, as cited in PATTON 2008: 1)*

In order to gain empirical insights into the current discourse and perspectives on systemic evaluation approaches of sustainable development in the case of German development cooperation, a qualitative research using two complementary streams was carried out. Firstly, semi-structured interviews among ten resource persons were conducted and processed through qualitative content analysis and PAIR (positions, actions, interests and roles) analysis. Secondly, four evaluation case questionnaires which provided real world evaluation examples were analysed and visualized by way of evaluation spectra and evaluation mapping. The purpose of this chapter is to highlight what the current evaluation discourse is about, in which ways sustainable development is actually conceptualized in evaluations and in how far systemic evaluation approaches are integrated therein. The latter includes questions on the perception about systemic approaches with regard to the perceived opportunities and limitations. All findings were analysed according to the study’s conceptual framework adapted from HUMMELBRUNNER (2013) and modified by differentiating actor-group clusters of ‘evaluation science’, including evaluation scientists and experts, ‘evaluation clients’, including evaluation managers and advisors, as well as ‘evaluation practitioners’, including freelance consultants and consulting companies. Each finding block is presented by a categorized summary table (as key result of the qualitative content analysis) and complemented by direct quotes to highlight the discourse as well as qualitative interpretations of the findings. At the end of this chapter, evaluation cases are depicted to visualize German development cooperation evaluations with regard to their contextual background and respective evaluation framing (shown in evaluation spectra and an evaluation mapping). To sensitively contextualize the findings and put them into perspective, a prior analysis on the positions, actions, interests and roles (PAIR-analysis) of the different actor groups was carried out.

2.1 Controversies: Positions, actions, interests and roles of different actor groups

By applying a ‘PAIR’ analysis, this section lays the basis for understanding how evaluation is differently perceived among the actor groups of evaluation scientists, clients and evaluators. PAIR analysis is a form of stakeholder analysis which aims to understand and characterize actors through their position, action, interests and roles on a certain stake (DIEPENMAAT 2011). In interpreting the PAIR matrix it is helpful to distinguish the top end of the matrix displaying the explicit ‘position’ and visible ‘action’ of the actor, from the bottom end of the matrix which depicts more invisible ‘interests’ and related ‘roles’ of actor (DIEPENMAAT 2011: 5).

Position	Action
<ul style="list-style-type: none"> - Evaluations need a scientific basis and analytical framework. - Systemic approaches can enhance the scientific basis and inform analytical frameworks. 	<ul style="list-style-type: none"> - Undertake basic and applied evaluation research; - Develop evaluation frameworks, designs concepts, methods and tools; - Undertake scientific evaluations; - Provide evaluation education and trainings; - Provide science based advisory and consultancy; - Undertake research on systemic evaluation approaches and disseminate findings; - Integrate systemic approaches and features in evaluation designs, education and trainings.
Actor: Evaluation Scientists & Expert	
<ul style="list-style-type: none"> - Promoting and enhancing evaluation science funding for evaluation science, evaluation sector; - Increase evaluation competencies and quality. 	<ul style="list-style-type: none"> - Developing evaluation theory to inform evaluation practice, advice and conducting evaluations.
Interests	Roles

Tab. 1: PAIR analysis of evaluation scientists & expert; Source: own analysis

The findings reveal that based on the interviewed evaluation scientists and expert’s² interest to enhance evaluation knowledge, competencies and quality, they take the role as ‘scientist and experts’ to generating viable knowledge on evaluation in theory and advice as well as engage in science-based evaluations in practice.

² Evaluation ‘expert’ is used in singular form as only one specialized evaluation expert was interviewed.

The respective position is that they promote a scientific foundation of evaluations which can also entail ‘systems theory’. It is noteworthy that though being scientists, they see their role also in advisory, consultancy and practice of evaluations. Characteristically, one expert noted:

“I also conduct evaluations and advice other evaluators and clients in designing evaluations” (EVALUATION EXPERT 2013).

The PAIR analysis of ‘evaluation clients’ shows that their main interest in evaluations is literally to gain ‘evidence’ for organisational/political decision making and use this evidence for justifying certain policies. It is their role and power to commission and consequently fund evaluations. The ‘utility’ of evaluations is pivotal to them and hence would also be a decisive factor in using systemic evaluation approaches. Through their actions, they basically frame and set the expectation level of evaluations.

Position	Action
<ul style="list-style-type: none"> - Utility of evaluation to policy and practice is pivotal. - Systemic approaches need to add value to evaluation utility, which is hardly the case yet. 	<ul style="list-style-type: none"> - Commissioning, framing, contracting, funding, steering, accepting, responding to and disseminating evaluations; - Quality standard setting and assurance of evaluations; - Testing systemic approaches for evaluation framing; - Dialogue and issue networking on systemic evaluation approaches.
Actor: Evaluation Clients	
<ul style="list-style-type: none"> - Evaluation as evidence, justification, input for policies, missions, goals, strategies, interventions, funding, support; - Representation and publicity. - Adhere to good governance. 	<ul style="list-style-type: none"> - Commissioning evaluations to meet corresponding evaluation demands by policy and practice.
Interests	Roles

Tab. 2: PAIR analysis of evaluation clients; Source: own analysis

Evaluators conduct evaluations as part of their (core) consultancy businesses. Hence, they are interested in competitively gaining clients and contracts, and satisfying them as well as their own methodological aspirations. From their point of view, evaluations are above all joint learning processes, whereby the evaluation focus is always determined by and/or negotiated with the clients’ expectations

and formalised through the contract and terms of references. Evaluators, hence, need to be flexible and have a comparative advantage to others. It is here where ‘systemic approaches’ can become attractive, however, only if it is demanded by the evaluation market, which so far has not been observed.

Position	Action
<ul style="list-style-type: none"> - Evaluation is above all a joint learning process in service of the client. - ‘Systemic’ is a new, trendy umbrella label with vague contents for evaluation. 	<ul style="list-style-type: none"> - Bidding for evaluation tenders; - Design, facilitate and report findings of evaluations within terms of reference; - Provide feedback and well justified recommendations; - Networking with other evaluators; - Integrating systemic features in evaluations.
Actor: Evaluators	
<ul style="list-style-type: none"> - Gain and satisfy clients; - Receive funding for evaluations; - Gain recognition and competitive advantage; - Have solid methodological and subject related standing; - Be qualified and competent; - Deliver high quality evaluations; - Contribute to learning and improvements. 	<ul style="list-style-type: none"> - Doing and coordinating evaluations and provide respective advisory to clients.
Interests	Roles

Tab. 3: PAIR analysis of evaluators; Source: own analysis

The PAIR analyses have shown that there are differing positions, actions, interests and roles with regard to evaluations in German development cooperation at large and systemic evaluation approaches in particular. The analysis, however, also brought to light that the different actor groups stand in specific relation to each other, for example, evaluation clients who commission and contract evaluations with evaluators and therefore, predominantly determine the purpose and way evaluations are designed and carried out; or for example, that roles and actions of certain actors groups overlap with others, such as evaluation scientists taking up the role of evaluators themselves and becoming advisors to evaluation clients. Based on the PAIR analysis, it now allows for a more nuanced examination of evaluation purpose, discourse, approaches and perceptions on systemic evaluation as differentiated by the actor groups.

2.2 Evaluation purpose: between learning and accountability

The purpose of evaluation is a core aspect of how evaluations are framed, designed and approached. It is hence also a critical part in understanding the opportunities and limitations of systemic evaluation approaches in relation to their appropriateness and accord with the overall evaluation purposes. This section thus provides a basic input for addressing the opportunities and limitations of systemic evaluation approaches.

<i>For what purposes are evaluations used in your organization?</i>	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - The main purpose is learning then accountability, controlling, legitimation, symbolisation function, publicity. - Evaluation contributes to capacity development, empowerment.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Main purpose is accountability then learning, knowledge production. - Evaluation contributes to improved strategic planning and action.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - The main purpose is institutional learning, accountability, publicity. - Evaluation contributes to quality management, knowledge exchange, strategic management and utilisation at project and regional level.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - The main purpose is learning then accountability, controlling and advisory. - Specific purpose is determined by client. - Evaluation contributes to capacity development, organisational development, strategic planning and participatory processes.

Tab. 4: Summary table on evaluation purpose; Source: own analysis

Most respondents claimed that evaluations are conducted for the purpose of learning and accountability. Learning is thereby, often equated with increased knowledge and insights on the interventions performance. A freelance consultant described these general goals like this:

“In terms of capacity development, we want to learn from evaluations, thus in the classical understanding of what went well and what less, and why. [...] And of course also accountability to the donor be it the BMZ or the World Bank. We also have to justify for what we have achieved and what not” (TRAINER, FREELANCE CONSULTANT 2013).

However, when looking at the priorities of evaluation aims, there are quite different positions on what an evaluation should fulfil. This can be seen between the different actor-groups but also within a specific group category.

For instance among the interviewed evaluators, there were claims that evaluations should merely “show how the projects are proceeding and how well the indicators and objectives have been achieved” (CONSULTANT 2013). Other evaluators oppositely stressed that “being able to start participatory processes through evaluations is most important [...] [i]t rather is the best when the findings are also obtained locally, and shared and discussed, and the partner organization then uses this and can further work with it” (FREELANCE CONSULTANT 2013).

Quite important for the clients group was also, however, that some degree of objective ‘evidence’ about the programme performance and results can be accounted for:

“What makes evaluations particularly valuable for me, is that you know more [...] on the basis of objective criteria to generate evidence so that you can say something is good or not so good” (EVALUATION MANAGER 2013),

or as mentioned by an NGO advisor:

“In the last 12 months we have had a development, which emphasizes more on accountability. This is due to the fact that we have to increasingly prove results” (NGO EVALUATION ADVISOR 2013).

Furthermore, the evaluation scientists and experts raised critical issues of evaluations uses, for instance to “publicly demonstrate a shiny image as hidden symbolisation function” (EVALUATION SCIENTIST 2013) and questioned whether learning from evaluations really takes place.

“The purpose of learning is indeed often desired but unfortunately cannot always be carried out in the scope that is envisioned and often stands in contradiction to justifications and control purposes, which an evaluation by nature also has” (EVALUATION EXPERT 2013).

In conclusion, the responses have highlighted that while there are certain generally accepted functions of evaluation (learning and accountability), they are differently prioritised among actors (accountability relatively higher among evaluation clients than among evaluators) who yet have different powers (with predominantly evaluation clients deciding about the specific purposes). In addition, there are various other utilization-related purposes, interests and hidden agendas (e. g. symbolisation) attached to evaluations, while others are apparently intended but practically hard to reach (e. g. joint learning). As will be discussed in chapter 3, systemic evaluation approaches seem to be more suitable for evaluations which explicitly focus on joint learning and less among control or accountability centred evaluations.

2.3 Current discourse on evaluations in the context of German development cooperation

The responses (see tab. 5) have shown that there are current prime discourses on (i) methodology such as randomized control trials (RCT) or (quasi) experimental designs, on (ii) structural issues such as the foundation of the new DEval institute or the fusion of GIZ and its related changes in the evaluation landscape of German development cooperation as well as holistic on (iii) generic trends in German development cooperation evaluations such as dealing with complexity, making best use of evaluations, and focussing on development results. The latter was responded by all actors group in unison, which shows its relevance for evaluation in German development cooperation. The following quotes highlight the main discourses:

Related to (i) methodologies: “The discussion in German development cooperation ranges from participatory methods at the grass root, for example, from the concept of the NGO IDEAS, up to the University of Saarland, CEval, Prof. Stockmann who exaggeratedly says that’s all ‘tea leaves reading’ [voodoo] and it is only about scientific approaches” (FREELANCE CONSULTANT 2013).

Related to (ii) structural issues: “The current situation, one can say, is characterized by a relative uncertainty caused by the formation of DEval by BMZ. Along with this uncertainty about the role of DEval – it is currently not clear in which direction it goes” (EVALUATION SCIENTIST 2013).

Related to (iii) generic trends: “Increasing emphasis is placed on results assessment [...] there is an end to only look on activities; we rather also look at their effects. That’s more than a trend – it’s already a standard” (FREELANCE CONSULTANT 2013).

What are the current discourses and trends on evaluation approaches and methodologies in your organization?	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - Founding of DEVAL and its role in GDC evaluation sector. - Scientific evaluation via rigorous impact assessment/randomised control trails (RCT) vs. participation and empowerment. - Evaluation under complexity and systemic evaluation approaches. - Results-based evaluation as steering instrument.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Discourses on feasibility of rigorous evaluation approaches and quasi- experimental designs, assessing efficiency, evaluability pre-conditions. - Trend to utilization focussed approach.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - Trend to increase evaluation utilisation and management though increased process evaluations. - Increased attention towards assessing results.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - Evaluation discourse context: shift from small projects to large programmes, more stakeholders, GIZ fusion and creation of DEVAL leading to evaluation standardisation, relevance for evaluations increased. - Evaluation under increased complexity. - Trend to more quantification, providing facts, rigorous evaluations using randomised control trials (RCT), comparative analysis. - Results orientation, focus on outcomes and impacts vs. logical framework. - Bridging grass-root participation and scientific approaches.

Tab. 5: Summary table on evaluation discourse and trends; *Source: own analysis*

Closely interlinked to the discourse, are the responses on ‘challenges’ faced in German development cooperation evaluations.

<i>Where do you see the greatest challenges for successful evaluations?</i>	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - High demands on evaluations in GDC vs. low resources and competencies. - Intransparency of evaluations and hidden agendas. - Finding appropriate methods for coping with complexity in impact assessments. - Strategic utilization of evaluations.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Political challenge to utilize evaluations in policy making. - Resistance on inconvenient evaluation findings. - Limited evaluability of interventions.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - Challenge to learn from evaluations in short- term emergency projects. - Ex-post evaluations difficult as no reference persons, logistics available. - Limited evaluation resources affect evaluation design and methods.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - Difficult to evaluate more complex interventions. - Balance scientific evaluation rigor with participation, pragmatism and practical limitations, holism vs. reductionism. - Balance conflicting purposes of controlling and learning. - Limited resources (time, budget, personnel) for stakeholder participation. - Challenges to reach objectivity, comparability of evaluations, track results. - Meaningful utilization of evaluations by stakeholders.

Tab. 6: Summary table on evaluation challenges; Source: own analysis

In addition to the methodological challenges of producing viable evaluations, and facing generic challenges, i. e. evaluations under conditions of complexity, moreover, practical considerations such as limited resources, low participation, and practical use of evaluations were outstanding responses. One typical example was provided by an NGO evaluation advisor:

“We have, because of limited resources, evaluations for periods of 2–3 weeks with field visits by external evaluators, which are still affordable. Along with that come clearly also methodological limitations, the whole discussion on rigorous comparative groups, RCT, etc., that is in our case relatively modest.

As for the evidence, one quickly enters into areas where we can work only with case-studies which prove quantitatively relatively little, or the cost would be incredibly high” (NGO EVALUATION ADVISOR 2013).

With regard to participation, a freelance consultant blatantly stated:

“What is to be evaluated is determined by donors. It would be nice if the local implementers could also have influence thereon because they also have certain issues” (FREELANCE CONSULTANT 2013).

This and the following statement indicate the underlying politics which frame and challenge the evaluations set up and use.

“An evaluation may show great results, may still be so viable, or inter-subjectively verifiable; should one not want to have the results, it is difficult to implement them. Depending on the subject, resistance was at times larger when the objects were too uncomfortable and less if the results were uncontroversial” (EVALUATION MANAGER 2013).

Looking at the contextual challenges in which evaluations are embedded, an evaluation expert re-stressed the difficulties related to conduct results-based evaluations in states of complexity.

“The biggest challenges for successful evaluations are, in my opinion, on dealing with complexity. And especially assessing results in complex situations. I notice that also with colleagues at conferences” (EVALUATION EXPERT 2013).

In conclusion, the discourses have confirmed the challenges which made systemic evaluation approaches relevant in response to coping with complex German development cooperation constellations as highlighted by HUMMELBRUNNER (2011, 2013) and focussing on development results (GIZ 2013). However, the dilemma in evaluation practice is that although addressing complexity is becoming a pivotal issue, evaluation resources are so restricted and the evaluability of interventions are often so limited that in many cases, no comprehensive systemic approaches can be conducted. Furthermore, do political interests and agendas ultimately play a key role in what type of evaluation is demanded and how it is actually utilised.

2.4 Evaluation of sustainable development

This section exclusively addresses how sustainable development is conceptualized in evaluation approaches among different actor groups in German development cooperation. The responses on evaluation approaches showed how sustainable development is understood and conceptualized in practice. Depending on the purpose of evaluation and adjusted to the local context, various types and approaches, as well as methods, were applied.

<i>How is the development cooperation impact on sustainable development and/or sustainability conceptualized in your evaluations?</i>	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	-Three dimensional analysis: - Horizontal- reconciling social, economic & ecological goals, - Vertical- global to local coordination, - Intergenerational justice. - For impact of capacity development: outcome mapping, outcome harvesting.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	- Sustainable development is vague, applied to everything, needs concretisation in strategies at various levels. - Sustained impact of interventions.
<i>Evaluation advisors</i> (non state actor) (n=1)	- Ex-post evaluations with mix of methods. - Focus on technical know how, sustained practices, institutional set-up, attribution to project.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	- Ex-post evaluations assessing sustainability of intervention results. - Using the three pillars of sustainability (economy, ecology, society) and intergenerational improvements. - Using sustainability indicators and guiding questions about capacities, competencies and where they stem from. - Acknowledging attribution problems and uncertainties.

Tab. 7: Summary table on sustainable development conceptualization in evaluations; *Source: own analysis*

Most commonly, sustainability is measured at ex-post level, hence a few years after the intervention has terminated. The understanding of sustainability is most often on ‘sustained impacts’ of interventions which is considered in relation to the three pillars of sustainability. Yet due to financial, logistical and personnel limitations, that rarely takes place. A freelance consultant critically noted:

“One can often only track sustained impacts several years after the project. And how many ex-post evaluations are there at all? I hardly know any, as there are hardly any financial resources for that kind of evaluation. Yet, exactly this would need to take place, if you want to investigate the effects of projects on sustainability” (FREELANCE CONSULTANT 2013).

In more details a consultant described evaluation questions related to sustainability indicators:

“[We used] a set of sustainability indicators in an ex-post evaluation. Evaluation questions related to sustainability are: How do you assess the capabilities of the partners to continue the project results? Are project results continued? Is even own money invested to continue working on the issues? These are questions that are based on OECD DAC, but also on various other institutions, which we then coordinate with the client of the evaluation” (CONSULTANT 2013).

When more openly asked about what evaluation approaches are applied, a wide variety of types, approaches and methods were named. General categorization was done according to temporal parameters (progress, mid-term, final and ex-post) and organizational responsibility (internal self and external evaluations). It is surprising that no respondent has stated the application of ‘ex-ante evaluations’. This may be related to the widely internalized OECD DAC definition of evaluations that concern ‘ongoing’ and ‘completed’ interventions³ in contrast to ‘appraisals’ focusing on the ex-ante assessments prior to funding.⁴

As being the most natural thing, one evaluation manager responded:

“Of course a mix of qualitative and quantitative methods is applied which, depending on the subject, often have to be combined. All evaluations have different approaches” (EVALUATION MANAGER 2013).

³ “The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results” (DAC 2009: 27).

⁴ “An overall assessment of the relevance, feasibility and potential sustainability of a development intervention prior to a decision of funding” (DAC 2009: 20).

What different evaluation approaches does your organization apply?	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - CEVAL approach for scientific evaluations. - Very wide spectrum. - For sustainable development evaluations: Process monitoring of impact.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Context specific mix of approaches and methods. - Results- based evaluation, contribution analysis. - Systemic evaluation framing.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - Internal, external evaluations. - Mid term, process, final, ex-post evaluations.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - Self evaluation, external evaluation, desk evaluations. - Process-, final-, ex-post evaluations. - Indicator frameworks, logical framework approach, results based monitoring and evaluation, participatory approaches. - Mix of approaches specific to terms of references and evaluation background.

Tab. 8: Summary table on evaluation approaches; *Source: own analysis*

However, a freelance consultant adds that the selection of approaches and methods is already framed by international and national evaluation standards, compounding German development cooperation, which are then requested by clients.

“The client decides [...]. But clients are tied to evaluation standards and these must then be satisfied because their hands are also bound, so to speak. I do not have much freedom there and I do not need to extensively discuss it with the clients as it is already fixed. It is also common to combine evaluations with the MDGs and the OECD DAC criteria. This influences the choice of methods and participatory approaches often suffer from that” (FREELANCE CONSULTANT 2013).

While all respondents expect that evaluations have certain uses for strategic and operational decision making, one evaluation manager addressed the issue of utilization with reference to MICHAEL PATTON (2008, 2011) and what that means for their evaluation approach.

“Generally one can say that we apply a utilization-focussed approach according to Patton in order to emphasize the learning aspect. Thus, how can the evaluation findings be utilised, by whom, how, in what form” (EVALUATION MANAGER 2013).

In conclusion to the research objective of exploring and clustering evaluation approaches of sustainable development, the findings have shown that in German development cooperation, practice of sustainable development evaluation is throughout the different actor groups mostly related to ex-post evaluations considering social, environmental and economic (plus in certain cases institutional or technical) aspects of ‘impact-sustainability’ at micro-project or programme level. This understanding of sustainability is similar to that of the CEval evaluation approach (MEYER 2007; STOCKMANN 2011). The conceptualization of sustainable development is hence mainly centred on the ‘impact’ and ‘sustainability’ of German development cooperation interventions in coherence with OECD DAC (2010a) criteria and relating to global policies such as the MDGs. Systemic evaluation approaches are thereby emerging in relation to framing sustainable development interventions (for example through results-based models) in respective evaluations or through overlapping elements with existing participatory evaluation approaches.

2.5 Systemic evaluation approaches: different understandings and connotations

This section discusses how systemic evaluation approaches are perceived by different actor groups and what opportunities and limitations they relate to systemic evaluation approaches. The findings reveal that systemic evaluation approaches are empirically not clearly defined. There are different understandings and connotations with the word ‘systemic’ among the respondents. While evaluation scientists draw their understanding from systems theory and recent publications of systemic evaluation approaches, evaluation clients and to a higher extent evaluators use approximations in explaining systemic evaluation approaches from their perception, experience and background. Nonetheless, the evaluators also inductively named key elements of systemic evaluation approaches which resemble the scientific discourse.

<i>How would you describe systemic evaluation approaches and what are the core characteristics?</i>	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - Generic concepts: interrelations, perspectives, boundaries. - Being systemic and thinking systemic. - Results modelling, from singular result chains to result systems models. - Features: system paradigm, social constructivism, social systems, reflexivity, complexity, diverse interest constellations, network focus, embeddedness, interactions.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Systems thinking and being systemic. - Generic concept of understanding interrelations, commitment about various perspectives and awareness on boundaries. - Applying a general, holistic view. - Not clearly defined. - What is not systemic? There is nothing purely mechanistic.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - Difficult to describe. - Features: Stakeholder perspectives, cause effect interrelations.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - I do not know; have not or marginally heard of it. - Focusing on diverse influencing factors and interrelations leading to results as conceptualized in results based monitoring. - Not linear but network based approach. - Characteristics of: system and system boundaries, system context, self- sustaining systems , system dynamics and uncertainty, cultural perspectives.

Tab. 9: Summary table on perceptions of systemic evaluation approaches; Source: own analysis

As a common element of systemic evaluation approaches, the idea of non-linearity was mentioned. According to a freelance consultant, systemic evaluation approaches meant to “consider instead of a linear model rather a network. There is much connected with each other and I know there is uncertainty and I cannot accurately capture it” (FREELANCE CONSULTANT 2013).

The notion of a network model has been also linked to GIZ as an important player in German development cooperation and related to their latest development of moving from a linear results-chain to an integrated results-model. An evaluation scientist stressed this development as such:

“GIZ has now made a decisive step and tries to withdraw from the singular results-chain to, in this sense, a more systemic model with feedback. Therefore, it is also incredibly complicated; but it perceives the matter more concrete and will lead to more modest impact claims” (EVALUATION SCIENTIST 2013).

Yet, stressing the complicated nature of defining ‘systemic’ evaluation approaches one respondent noted:

“With the term systemic I associate an overall context relationship where things are in inter-relationships und where one has to see something as a whole and where particularistic perspectives are too limited; [...] probably the term is not defined at all” (EVALUATION MANAGER 2013).

Even more critical, a freelance consultant stressed the abstract and theoretical nature of ‘systems’ approaches in theory which, when it translates into practice, does not mean anything new to what is already practiced:

“But the whole [systems] thing is a very theoretical model to me because the ideas behind it are not new and what it really means in practice, to consider different perspectives is nothing new” (FREELANCE CONSULTANT 2013).

Further focusing on the evaluation practice in German development cooperation, most respondents stated that they do partly apply systemic evaluation approaches, although they do not explicitly refer to them as that.

Respondents claimed that there is a clear overlap with participatory approaches, and results-based evaluations which are both very prominent in German development cooperation. By providing insights into his evaluation practice one freelance consultant stated:

“Systemic approaches are difficult to break down in practice. In practice, I partly use it [...] for example, to look at a situation from different perspectives and also ask questions accordingly. These are hypothetical questions, ‘imagine you are ... are in the position ... how would you assess that?’ Then I ask the NGO worker ‘how would a government official see it?’ I also make a so-called re-framing, the classic example, when someone tells me the glass is half full, then I ask ‘could you not see it as half empty or conversely?’ This has much to do with participation and that is, that I look at things from the perspective and with the knowledge of the target groups” (FREELANCE CONSULTANT 2013).

<i>Are systemic evaluations approaches conceptualized and/ or applied in your organization? If so how?</i>	
Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - Yes, outcome mapping, outcome harvesting, process monitoring of impact, circular questioning, systemic questioning, reflecting team, re-framing. - Partly, already integrated in existing approaches (e.g. CEVAL approach).
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	- Systemic analytical frameworks for evaluations.
<i>Evaluation advisors</i> (non state actor) (n=1)	- Partly, via analysis of stakeholders, problem cause-effect interrelations.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - No, I did not apply, so called, systemic approaches. - It was never demanded to apply systemic approaches in evaluations. - Yes: results models and analysing interrelations, systemic questioning tools, reframing from different perspectives, dialogue on system boundaries in planning, integrating perspectives via stakeholder participation at different levels, stakeholder analysis, systems tools of capacity works toolkit. - As evaluator I use basic systems thinking.

Tab. 10: Summary table on implementations of systemic evaluation approaches; *Source: own analysis*

However, also prior to the actual inquiry phase, systemic approaches help to frame evaluations and clarify ‘stakes’ of stakeholders in a given evaluation. An evaluation manager accordingly explained:

“In a workshop [...] the evaluation framing had to be developed [...] namely, to define exactly what is my purpose and which stakeholders are involved and what stakes do they have. So the distinction between the stakeholders and their stakes and get that clear and only then derive the criteria and objectives of the various specific evaluation questions, which then should answer the impact” (EVALUATION MANAGER 2013).

There were furthermore clear affirmations that systemic evaluation approaches are used, for instance, within the wider management framework of GIZ. A freelance consultant claimed:

“We do this and it is termed accordingly! Due to the GIZ management model of capacity works, it is implemented. Because it focuses strongly on that; there are tools about stakeholder analysis, cooperation landscape, and also to system boundaries in a first step” (TRAINER, FREELANCE CONSULTANT 2013).

In contrast, the opposite was also mentioned. Another freelance consultant manifested:

“I was never asked, do your evaluation systemically” (FREELANCE CONSULTANT 2013).

Besides, it remains debatable what is actually considered ‘systemic’ and what is not. Being asked about whether systemic evaluation approaches are applied in their organization, an evaluation manager returned the question and prompted:

“That you can tell me; if that what you see in the evaluation reports and methodology annexes are considered systemic approaches and methods in your opinion or not” (EVALUATION MANAGER 2013).

From these responses certain positions towards systemic evaluation approaches already emerged, which in the following question were then made more explicit.

As shown in table 11, the positions ranged from positive attitudes of considering systemic evaluation approaches as an added value for the evaluation portfolio, especially in complex situations, to more negative and sceptical views. It shows that there is certain confusion around the understanding of systemic evaluation approaches and hence the hesitant up to resistant reactions. As the latter dominated the responses, a few are listed in the following:

“Regarding systemic approaches, there is still insufficient experience among us. We are working on it, we also see the need, but we further come in vast methodological difficulties, if you would approach it scientifically. There is much anecdotal, self-tailoring in it” (NGO EVALUATION ADVISOR 2013).

“They all say that one should look at the structure and context. I would know anything where you would not say that. If you mean that with systemically, I find it obvious. That’s why I was a bit suspicious of what they mean by the term. So [something] purely mechanistic, I would not know at all” (EVALUATION MANAGER 2013).

“That existed before but was only named differently. So I think, ok if that is systemic, then fine, but we have designed and used it already. Either there is not so much new to it or I do have to deal with it even more” (FREELANCE CONSULTANT 2013).

**What is your position on systemic evaluation approaches?
Where do you see strength (advantages), weaknesses (disadvantages), opportunities or threats do you see with regard to systemic evaluation approaches for your organization?**

Group/Actor:	Categorized answer summaries
Evaluation Scientists <i>Evaluation scientists, experts</i> (n=3)	<ul style="list-style-type: none"> - Systemic approaches enhance method spectrum of evaluations. - Strength: Complexity, uncertainty and interrelations are better addressed and understood in systemic approaches. - Weakness: Too abstract and complicated, requires high abstract thinking, not sensitive to different cultures, difficult to produce clear positions. - Opportunities: Better understanding of real world complexity in programme planning, enhanced learning potential, improved linkage to systemic advisory. - Threat: Isolation and schism with 'non systemic' people, critique nothing new.
Evaluation Clients <i>Evaluation managers</i> (state actor) (n=2)	<ul style="list-style-type: none"> - Suspicious about the term 'systemic'. - Systems thinking and being systemic is helpful in evaluation framing. - Strength: Systemic perspective structures planning, design of evaluations. - Weakness: Confusion with participatory approaches, resistance and schism of systemic and non systemic.
<i>Evaluation advisors</i> (non state actor) (n=1)	<ul style="list-style-type: none"> - Position: hesitant, insufficient know-how, expect methodological difficulties, seems not ripe yet. - Strength: Improve understanding of interrelations, uncertainties, enhance partner dialogue, widen the perspectives frame. - Weaknesses: Feasibility of systemic approaches, coping with complexity. - Threat: Difficult to determine what is systemic and what not.
Evaluators <i>Consultants</i> (freelance, consortia) (n=4)	<ul style="list-style-type: none"> - Do not understand what is new; in reality not new but only new label. - It will never become a general approach. - Strength: more objective, more precise for learning, more transparency but perspective dependent, enhances stakeholder participation. - Weakness: too expensive, complicated, difficult, not feasible; uncertainty not wanted by donors, clients. - Threat: organisations become susceptible to negative evaluation findings - Opportunity: DEGEVAL association has a working group on systemic approaches.

Tab. 11: Summary table on positions of systemic evaluation approaches; Source: own analysis

It can be concluded that systemic evaluation approaches are often perceived as abstract and unclear among respondents. They are met with a certain resistance which stems from the absence of experience among evaluation practitioners as well as clients and moreover the confusion of what makes an evaluation systemic, and whether it is truly new or old but apparently in a new disguise. While evaluators in German development cooperation have certain elements of systemic approaches already included in their evaluations, they however relate them to existing participatory approaches or results-based evaluations. With regard to opportunities, all actor groups relate systemic evaluation approaches with gaining better insights into the complex reality of development interventions. The perceived limitations of complicatedness and vagueness, being too theoretical and less practical, being too time consuming and costly are however still dominating and enforcing latent resistance. As such evaluation practitioners and moreover clients are so far hesitant to embark on 'more' systemic evaluation approaches.

2.6 Taking a closer look: Evaluation cases from Bolivia, the Philippines, Nigeria and Ethiopia

The evaluation cases provide real world examples of evaluation in German development cooperation. Through a sample of four evaluation cases it examines the context specific application of evaluation framing vis-à-vis the intervention background.⁵ These examples allow a micro-perspective on what empirically was evaluated, using which methods, considering what aspects and including which perspectives within the respective problem and intervention context.⁶ This section, hence, provides evaluation practice insights on the conceptualization of sustainable development evaluations and the application of systemic approaches in sustainable development evaluations. For this purpose, the findings of four sample evaluation examples from Bolivia, the Philippines, Nigeria and Ethiopia are visualized through evaluation spectra (see fig. 2). The evaluation spectra show and measure the characteristics of the intervention problem background, components and stakeholder groups vis-à-vis the stakeholders involved in the evaluation, the applied evaluation methods, perspectives considered in the evaluation, sustainable development dimensions and specific issues covered in the evaluation. A full explanation of the meaning of each category is presented in the evaluation spectrum legend (see fig. 2).

⁵ As of confidentiality, the implementing organisation as well as the actual intervention is not named. Instead, the country name, intervention sector and intervention level are depicted.

⁶ The presented evaluation cases are based on specific evaluation case questionnaires.

Subsequently, a mapping of these cases is undertaken along an axis framework of ‘simple’, ‘complicated’ and ‘complex’ intervention framings on the x-axis and the degree of systemic elements used in the evaluation framing on the y-axis (see fig. 3). While the intervention frame was attributed to a direct corresponding question in the evaluation case questionnaire, the classification of the systemic features derived from a set of five questions. The selection of the five questions was done according to the core features of systemic evaluation approaches as stipulated by WILLIAMS and HUMMELBRUNNER (2011), encompassing multiple ‘perspectives’ (relating to the ‘evaluation perspective’ question), ‘interrelationships’ (relating to the questions of ‘evaluation dimensions’ and ‘evaluation issues’) and ‘boundaries’ (relating to the ‘evaluation stakeholder groups’ question) as well as ‘thinking systemically’ (relating to ‘evaluation methods’).⁷ All checked answers were then summed up and compared to the maximum possible affirmed answers.⁸ The vertical positioning along the ‘low and high continuum’ of systemic features in evaluation hence represents the relational degree of the specific evaluation case in comparison to the maximum (high) or minimum (low) possible value. The underlying classification system is presented in table 13.

In the following, the four evaluation cases are briefly described, illustrated and compared to each other.

The evaluation case of rural community development in Bolivia illustrates an example of a regional German development cooperation intervention which provides technical advice, infrastructure development, capacity development and multi-actor networking (four components) to address social, economic, environmental and institutional problems (four problem dimensions) and involves more than ten stakeholder groups in the intervention. Against this background, the respondent classified the intervention context as ‘complex’. To evaluate the intervention’s impact on sustainable development, an ex-post impact evaluation, aimed primarily at ‘learning’ was jointly developed by donors, implementer, partners and the consultant (four evaluation stakeholder groups). Pursuing a results-based framework, the evaluation used five different methods, being indicators sets, outcome mapping, participatory tools, circular dialogue and kids inquiry games. As a participatory evaluation, it included in the inquiry the perspectives of donors, implementers, partners, stakeholders and target groups (five evaluation perspectives). The intervention results were assessed with regard to social/cultural, economic, environmental and institutional dimensions (four evaluation dimensions) and also specifically included unintended results, various knowledge bases, the intervention system, and sustainability. Based on the way the evaluation was carried out, i. e. the evaluation stakeholder groups, the evaluation methods, the integrated evaluation perspectives, the evaluation dimensions and issues

⁷ These features are subsequently referred to as systemic elements or features of systemic evaluation approaches.

⁸ Accordingly, it is to be interpreted as: the more checked answers, the higher the value and the more systemic features in evaluation framings.

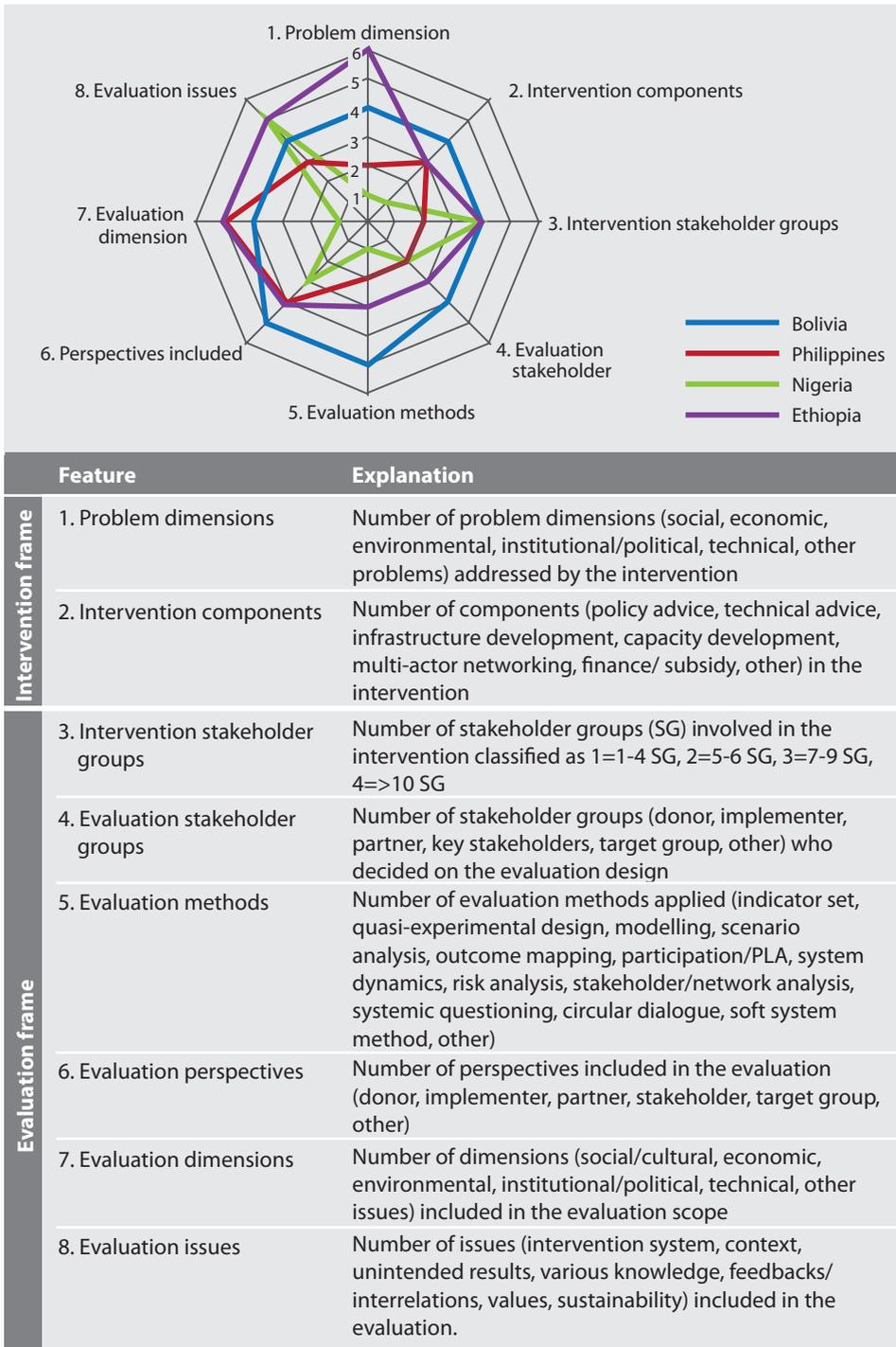


Fig. 2: Compilation of evaluation spectra; Source: own analysis

assessed, the evaluation framing displayed 22 out of 39 possible ‘systemic’ features (see tab. 13). In figure 3 it eventually illustrates that, while having a ‘complex’ intervention framing, the evaluation applies an assessment with ‘medium-level’ systemic features which mainly stem from its participatory and results-based evaluation design.

	Bolivia	Phillipines	Nigeria	Ethiopia
1. Problem dimensions	4	2	1	6
2. Intervention components	4	3	1	3
3. Intervention stakeholder groups	4	2	4	4
4. Evaluation stakeholder groups	4	2	2	3
5. Evaluation methods	5	2	1	3
6. Evaluation perspectives	5	4	3	4
7. Evaluation dimensions	4	5	1	5
8. Evaluation issues	4	3	5	5

Tab. 12: Results from evaluation cases; Source: own analysis

The second evaluation case of a food security intervention in the Philippines took place at district level where it mainly addressed social and food security problems (two problem dimensions) through providing technical advice, capacity development and multi-actor networking (three intervention components). While the respondent stated that 5–6 stakeholder groups were involved in the intervention, the overall intervention context was perceived as ‘complicated’. The commissioned process evaluation which was designed exclusively by the donor and implementer (two evaluation stakeholder groups) aimed at contributing to learning, stakeholder empowerment and expert judgment. Based on the intervention’s logical framework the evaluation applied participatory rural appraisal (PRA) methods and indicator sets of the logical framework matrix (two evaluation methods) to assess five evaluation dimensions (social/cultural, economic, environmental, institutional & technical dimensions) focusing at three systems issues (intervention system, unintended results, feedbacks & interrelations). Due to its participatory design, the evaluation included the perspectives of the donor, partner organization, stakeholders and target group (four evaluation perspectives). With regard to the evaluation framing the evaluation incorporated a few distinct systemic features mainly due to its participatory approach and hence was classified as having ‘lower medium’ systemic features (due to applying 16 out of 39 systemic features, see tab. 13).

The subsequent case of ‘slum development’ in Nigeria exclusively focussed on capacity development (one intervention component) to tackle institutional problems (one problem dimension) at district level. As a result of these characteristics, the respondent viewed the intervention context as ‘simple’. The process

evaluation which assessed the intervention was designed by the donor and the partner organization (two evaluation stakeholders) for the purpose of expert judgment, controlling and learning. Based on the logical framework matrix of the intervention, an indicator set (one evaluation method) was used for the evaluation of institutional aspects (one evaluation dimension) with regard to the intervention system, context, unintended result, various knowledge, feedbacks/interrelations and sustainability (five evaluation issues). Within the assessment, the evaluation included the perspectives of the implementation partner, stakeholders and the target group (three perspectives). On the whole, the evaluation framework did include minor systemic features as a result of the single method pursued and dimension assessed (with 12 out of 39 systemic features, classified as 'low'; see tab. 13).

Within the last case of water supply and sanitation in Ethiopia, a final programme evaluation was carried out to render strategic learning and accountability. The intervention addressed multiple problems (social, economic, environmental, institutional, technical and water as well as sanitation access) by primarily providing technical advice, infrastructure development and capacity building (three intervention components) in an integrated watershed management approach. As a result of this technically difficult approach involving a vast group of stakeholders, the respondent classified the intervention context as 'complicated'. According to the demand of the donor, implementer and local partner organization (three evaluation stakeholders), the corresponding evaluation used the programme's logical framework matrix and OECD DAC evaluation criteria for a multi-dimensional assessment of programme results (social/cultural, economic, environmental, institutional, technical dimensions). By using an indicator set and a pragmatic quasi-experimental design, two evaluation methods were applied to interact with implementers, partners, stakeholders and the target group (four evaluation stakeholders) by focussing on the intervention system, context, (un-)intended results, local knowledge and sustainability issues (five evaluation issues). While having a 'complicated' intervention framing, the evaluation framing was classified due to its substantial multi-perspectives, participation and issue diversity as having 'medium' systemic features (20 out of 39 systemic features, see tab. 13).

Figure 3 illustrates all four evaluation cases in relation to their perceived (simple, complicated or complex) intervention framing and (low to high systemic) evaluation framing.⁹

9 Note: The axis framework was designed using Microsoft excel's scatter plot function. The positioning of evaluation cases along the vertical axis of 'evaluation framing' was based on a continuum between the minimum value of 5 representing 'low' and the maximum value 39 representing 'high' systemic features. However, in critical reflection of the subjective nature of this mapping, it was decided not to show the actual values (between 5 & 39) on the y-axis to prevent claiming pseudo quantification and measurement exactness of 'systemic' features. It is by no means aimed to 'measure the immeasurable', but rather to stimulate qualitative discussions on the relation of systemic evaluation approaches in relation to the complexity of interventions.

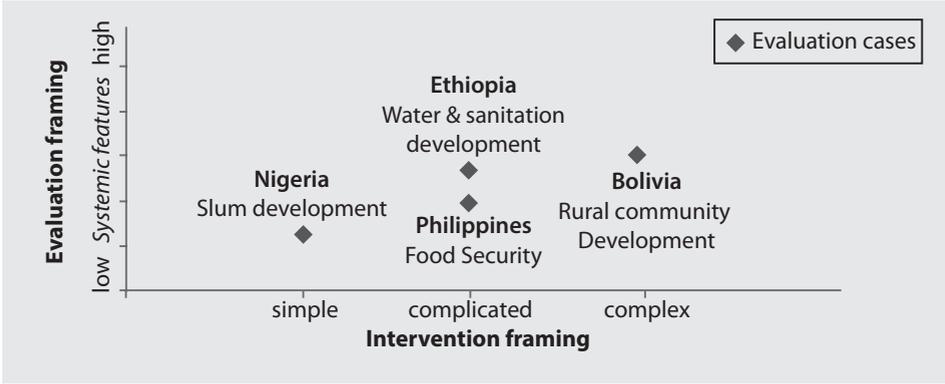


Fig. 3: Evaluation case mapping: intervention vs. evaluation framings; Source: own analysis

While such a small sample is certainly not sufficient to draw viable or representative conclusions, the mapping of empirical evaluation cases, however indicate that within this study the interventions showing a higher level of complexity also applied more systemic features within the respective evaluation. While the slum development intervention in Nigeria took place in an allegedly ‘simple’ context, the evaluation approach also did not apply substantial systemic features. In contrast, the Bolivian rural community development intervention was marked as ‘complex’ and as a result also applied more systemic features in the evaluation to explore, sense and interact among the various evaluation stakeholders. The cases of the Philippines and Ethiopia emphasize that although having comparable generic framings, it furthermore depends on the specific circumstances, stakeholder interests and aims related to the evaluation to what extent systemic features are taken up in the evaluation approach. The following table 13 summarizes all underlying values to the evaluation case mapping of Figure 3.

GDC Interventions		Bolivia	Phillipines	Nigeria	Ethiopia	
Values Max Min		Rural community development	Food security	Slum development	Water & sanitation development	
Intervention framing		complex	complicated	simple	complicated	
Systemic features in evaluation framing		medium	lower medium	low	medium	
Systemic feature	Evaluation stakeholder groups	6 1	4	2	2	3
	Evaluation methods	13 1	5	2	1	3
	Evaluation perspectives	6 1	5	4	3	4
	Evaluation dimensions	6 1	4	5	1	5
	Evaluation issues	8 1	4	3	5	5
Total		39 5	22	16	12	20

Tab. 13: Sample evaluation cases; Source: own analysis

In conclusion, except in one case, all evaluations have considered the three pillars of sustainability or a 'triple bottom line' (TBL) concept, plus the institutional and/or technical aspects of sustainability at micro-intervention level. Evaluations were typically designed by the donor in conjunction with the implementing agency and to a lesser extent also involved the partner organization in the developing country. A distinct systemic feature which was widely integrated in the evaluation cases was the use of various perspectives through participatory approaches. Furthermore, most evaluations considered interrelations, the interventions system and its context as well as emerging results and unintended effects in their inquiries. Looking at the pattern of the sample evaluation cases, it can be observed that interventions with higher complexity also had more systemic elements in their evaluation. However, while the findings have shown that elements of systemic evaluation approaches are indeed present in German development cooperation evaluations, they nevertheless lack a systemic evaluation framework which logically combines the different elements into sound and flexible approaches.

3 DISCUSSION – WHAT IS NOT SYSTEMIC?

Following the findings of the research, it remains to be answered what the results mean with regard to the opportunities and limitations of systemic approaches in sustainable development evaluations of German development cooperation.

OPPORTUNITIES AND LIMITATIONS OF SYSTEMIC EVALUATION APPROACHES IN GERMAN DEVELOPMENT COOPERATION

According to the evaluation science discourse on systemic evaluation approaches in sustainable development (BAWDEN 2007; HUMMELBRUNNER 2007, 2013; MIDGLEY 2000; PATTON 2011; REYNOLDS, WILLIAMS 2012; WEAVER, ROTMANS 2006; WEAVER et al. 2007; WILLIAMS, HUMMELBRUNNER 2011; WILLIAMS, IMAM 2007), systems approaches provide distinct solutions for evaluations in complex environments. Among the solutions are developing appropriate evaluation frameworks, sensitively capturing complex situations, holistically analysing sustainable development programmes or policies within their respective environment and fostering mutual social learning for adaptive management. There is a paradigm evolving which holds that combining systems thinking with evaluation into integrated systemic evaluation approaches can better reflect the complexity and multi-dimensionality of sustainable development (HUMMELBRUNNER et al. in this book; ROTMANS 2006; WEAVER, ROTMANS 2006; WILLIAMS, HUMMELBRUNNER 2011). Yet, so far, not much empirical research on applications of systemic evaluation approaches has been carried out, particularly not in the case of German development cooperation.

Hence, by drawing from the findings of this research, it seems that systemic evaluation approaches are not widely applied in German development cooperation, at least not under the term ‘systems/systemic’. However, certain elements thereof are indeed integrated into evaluations. The comparative analysis of four evaluation cases has found that a substantial use of systemic elements in evaluations was prevalent in interventions which themselves are framed with high complexity (see fig. 3)¹⁰. Nevertheless, most respondents from the evaluation client and practitioner actor groups have reservations towards systemic evaluation approaches

¹⁰ Although the patterns in figure 3 would suggest a positive relation between increased intervention complexity and increased systemic elements in evaluations, this is purposefully not claimed in this study due to the small sample size.

which stem from the perceived abstract nature of systems concepts, absence of experience and confusions with other evaluation approaches, in particular, participatory approaches. While most respondents on the one side assume distinct potentials of systemic evaluation approaches in complex environments, on the other side they relate various practical barriers and conflicts of interest with them.

The empirical findings show that actors from within the clients' domain mostly relate the opportunities of systemic evaluation approaches to improved framing and the designing of evaluations under complex conditions, fostering dialogue with stakeholders and improving understanding of interrelations and uncertainties or risks. Evaluators conversely stressed the enhanced repertoire of the evaluation inquiry, the improved learning potential and better engaging stakeholder participation. A distinct opportunity is furthermore seen in the German Evaluation Society's (DeGEval) working group on systemic approaches which provides a forum for discourse, negotiation and standard setting.

These findings are to a large extent in line with and/or complementary to HUMMELBRUNNER (2013: 34–35) who claims four main potentials of systemic evaluation approaches, namely, (i) extending the evaluation approach portfolio for complex 'evaluands', (ii) appropriate simplification of complex situations, (iii) enhanced assessment of intended systems change and (iv) general sense-making in complex situations.

With regard to the perceived limitations of systemic evaluation approaches, it was (surprisingly) stated even from the evaluation science respondents of this study that systemic evaluation approaches are too abstract, complicated, might not be sensitive to different cultures, and eventually be difficult to produce clear findings and recommendations that can be used and are required by evaluation clients. Evaluation clients in turn see the main limitations in the appropriateness, feasibility and cost-effectiveness of systemic evaluation approaches with regard to its actual use in relation to enhancing political accountability and legitimacy. Evaluators, eventually, criticise the confusion of systemic evaluation approaches with already existing participatory approaches, while risking a substantial schism in the evaluation field between systemic and non-systemic evaluations. They further imagine systemic evaluation approaches to be too expensive, complicated as well as difficult and while making uncertainty more explicit, such approaches presumably conflict with the expectation by donors and clients of evaluations.

These perceived limitations stand in contrast to the ascribed potentials of systemic evaluation approaches claimed by HUMMELBRUNNER (2013), PATTON (2011) as well as WILLIAMS and IMAM (2007). However, HUMMELBRUNNER (2013) acknowledges that certain frame-conditions can be considered as favourable or unfavourable to integrating systems approaches in evaluations. While an evaluation focus on mutual learning, participation, experimentation and reflection is considered as conducive to systemic evaluation approaches. HUMMELBRUNNER (2011: 35) holds that an understanding of evaluation as predominantly 'objec-

tive' valuing or as an external control mechanism with strong political-economic interests, is detrimental to the application of systemic evaluation approaches.¹¹ It hence re-emphasizes the critical aspects of "fitness for purpose" (WEAVER 2013) and "salience, credibility and legitimacy" (CASH et al. 2003: 8086) which systemic evaluation approaches need to fulfil as they feed back and compete for attention in the political decision making processes (WEISS 1993: 94).

Hence, by considering these opportunities and limitations of systemic evaluation approaches, it might explain why 'in theory', many actors in German development cooperation are interested in systemic evaluation approaches and appreciate the potential added value while 'in practice', they still have not introduced systemic evaluation approaches on a significant scale but rather in certain parts according to the specific intervention context, interests and evaluation purposes at hand.

CRITICAL REFLECTION

"What is not systemic? [...] there is nothing purely mechanistic!"
(EVALUATION MANAGER 2013).

This statement by an evaluation manager ignites a critical reflection on what actually can be called 'systemic' evaluation approaches and what can not. While in theory and respective models it is simple to draw simplistic, mono-causal or linear cause effect relationships, in reality it shows that this is not the case. For the most part this has hardly ever been assumed in German development cooperation planning and evaluation as claimed by the EVALUATION MANAGER (2013). Hence, in practice, it is more appropriate to depict a continuum between systemic and non systemic elements in evaluations as indicated in figure 3: Evaluation case mapping intervention vs. evaluation framings. However, the logical question arising would be, when is an evaluation approach systemic and when not?¹² Or rather, at which threshold is an evaluation aptly called 'systemic' and who decides what can be called systemic and what not? As there are furthermore certain overlaps to existing evaluation approaches such as 'participatory approaches' or 'results-based approaches' it furthermore creates a certain amorphousness which has led to characteristic statements such as "that existed before but was only named differently" by a FREELANCE CONSULTANT (2013) or "there is much anecdotal, self-tailor made in it" by an NGO EVALUATION ADVISOR (2013).

Even from the side of theory, HUMMELBRUNNER (2013: 5) acknowledges that there is no clear differentiation possible between 'systemic evaluation' and 'non systemic evaluation' which actually is neither possible nor useful to undertake.

¹¹ This furthermore holds true for evaluations with low flexibility as of overly restricted terms of references, time and/or resources (HUMMELBRUNNER et al. in this book).

¹² Especially the criterion of 'being systemic' leaves ample room for debate.

HUMMELBRUNNER (2013: 5) further explains that as systems thinking is not based on a homogenous and complete theoretical framework, the term ‘systemic’ is not finally defined and hence its diverse connotations are in part already known in German development cooperation but applied under different labels. In this context, a FREELANCE CONSULTANT (2013) exemplarily stated:

“I partly use it [systemic evaluation approaches] [...] for example, to look at a situation from different perspectives. [...] This has much to do with participation” (FREELANCE CONSULTANT 2013).

HUMMELBRUNNER (2013) nonetheless argues that such vagueness should not lead to abandoning the ‘systems’ label, nor should it be overtly used as trendy term for each and everything related to systemic elements. What can be interpreted from the findings of this study is however, that the core idea of “thinking systemically and being systemic” in evaluations (BAWDEN 2007; HUMMELBRUNNER et al. in this book; WILLIAMS, HUMMELBRUNNER 2011) although sporadically emerging, has not yet crystallized into widely agreed concepts, understandings and actions. The latter can be claimed for all three actor groups of evaluation science, clients and practitioners as assessed in this research.

4 CONCLUSION

Set amidst the transdisciplines of sustainability science and evaluation science, this article discussed systemic approaches in sustainability evaluations in the case of German development cooperation. While aiming to contribute towards sustainable development, German development cooperation interventions increasingly depend on interactions among multi-stakeholders at multiple-levels, encompassing multi-faceted measures. Against this backdrop, the need for systemic evaluation approaches has emerged to more accurately capture and deal with the increasing complexity faced by German development cooperation. A review of evaluation theory has shown that there is not one single 'systemic evaluation' approach but a diversity of systemic concepts and approaches used for evaluations under complex conditions. The common basis of the various systems approaches is however their focus on multiple 'perspectives', dynamic 'interrelations' and 'boundary' judgements. Besides the notions of 'thinking systemically' and 'being systemic', the framework of developmental evaluations is emerging as a key characteristic of systemic evaluation approaches. From recent developments in sustainability science, 'integrated sustainability assessment' delivers a promising systemic framework that aims to capture sustainable development in its complexity and hence can be appropriate also for the case of German development cooperation.

The empirical findings of this study have shown that while the relevance of systemic evaluation approaches in response to coping with complex German development cooperation constellations is widely accepted, the dominant evaluation practices pursue rather mechanistic evaluation approaches. Comprehensive systemic evaluation approaches are hardly practiced as a consequence of low evaluation resources (financial, time and qualified human resources), limited evaluability of interventions and widespread hesitation of evaluation clients and practitioners due to its perceived impracticability and vagueness. At present, sustainable development evaluations are mostly related to ex-post evaluations considering social, environmental and economic (plus in certain cases institutional or technical) aspects of sustainability at micro-project or programme level. The conceptualization is hence mainly related to sustainable impacts of German development cooperation interventions in coherence with OECD DAC criteria and relating to global policies such as the MDGs.

Identified elements of systemic approaches in evaluation practice are thereby related to framing sustainable development interventions through results-based models and existing participatory evaluation approaches. Most common is thereby

the integration of various perspectives in the inquiry of evaluations, considering interrelations within the interventions system and its context, as well as tracking emerging results and unintended effects. However, while the findings have shown that elements of systemic evaluation approaches are indeed emerging in German development cooperation evaluations, in particular in interventions with higher complexity, they nevertheless mostly lack a systemic evaluation framework which combines the different elements, methods and tools into sound and flexible approaches.

With regard to the perceived opportunities of systemic evaluation approaches in the context of German development cooperation, gaining better insights into the complex reality of development interventions, extending the repertoire of evaluation inquiry, improved learning and improved ways of engaging stakeholders into participation processes was found. A distinct opportunity is furthermore seen in the German Evaluation Society's (DeGEval) active working group on systemic evaluation approaches which functions as a loose issue network to exchange knowledge and foster discourse. The perceived limitations of complicatedness and vagueness, being too theoretical and less practical, being too time consuming and costly are however still dominating and enforcing hesitation. As such evaluation practitioners and moreover clients are, so far, hesitant to embark on 'more' systemic evaluation approaches. By considering the framework and purpose of evaluation, it appears that systemic evaluation approaches are eventually more suitable for 'open' evaluations which explicitly focus on 'joint learning' and/or 'participation' and less among evaluations centred on 'control' or 'accountability'.

REFERENCES

- ARBTER, K. (2007): Strategic Environmental Assessment and Sustainability Impact Assessment: two participatory assessment tools for sustainability. In: Störmer, E.; Schubert, U. (Eds.), *Sustainable development in Europe: concepts, evaluation and applications*. Cheltenham: Edward Elgar Publishing, pp. 83–95.
- BÄCKSTRAND, K.; KHAN, J.; KRONSELL, A.; LÖVBRAND, E. (2010): The promise of new modes of environmental governance. In: Bäckstrand, K.; Khan, J.; Kronsell, A.; Lövbrand, E. (Eds.), *Environmental Politics and Deliberative Democracy. Examining the promise of new modes of governance*. Cheltenham: Edward Elgar Publishing, pp. 3–27.
- BAWDEN, R. (2007): A systemic evaluation of an agricultural development: A focus on the worldview challenge. In: Williams, B.; Imam, I. (Eds.), *Systems concepts in evaluation*. Point Reyes, CA: Edge Press, pp. 35–46.
- BOND, A.; MORRISON-SAUNDERS, A. (2011): Re-evaluating sustainability assessment: aligning the vision and the practice. In: *Environmental impact assessment review*, 31(1) doi: <http://dx.doi.org/10.1016/j.eiar.2010.01.007>, pp. 1–7.
- CASH, D. W.; CLARK, W. C.; ALCOCK, F.; DICKSON, N. M.; ECKLEY, N.; GUSTON, D. H.; MITCHELL, R. B. (2003): Knowledge systems for sustainable development. In: *Proceedings of the National Academy of Sciences*, 100(14), pp. 8086–8091.
- DAC (2010a): *Evaluating development co-operation. Summary of key norms and standards (Second ed.)*. Paris: Organisation for Economic Co-operation and Development (OECD).
- DE RIDDER, W.; TURNPENNY, J.; NILSSON, M.; RAGGAMBY, A. VON (2007): A framework for tool selection and use in integrated assessment for sustainable development. In: *Journal of environmental assessment policy and management*, 9(04). doi: 10.1142/S1464333207002883, pp. 423–441.
- DEVAL (2013): *Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit*. from <http://www.deval.org/de/about-us.html>. Retrieved 23.07.2013,

- DIEPENMAAT, H. (2011): Understanding actors in isolation: the PAIR-analysis. Chapter 12 (translated from Dutch; complete translation in progress). In: H. Diepenmaat (Ed.), *Multi-actor process management in theory and practice*. Almere: Parthenon Publishers, pp. 1–18.
- GIZ (2012): *Das Wirkungsorientierte Monitoring in der GIZ. Orientierungsrahmen*. Eschborn: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.
- GIZ (2013): *Measuring – Assessing – Making improvements. Findings and conclusions from monitoring and evaluation, 2010–2012*. D. Giesen-Thole (Ed.) Retrieved from <http://www.giz.de/de/downloads/giz2013-en-evaluierungsbericht-2010-2012.pdf>
- HARDI, P.; MERTINUZZI, A. (2007): Series editors' preface: Evaluating Sustainable Development – topics, trends and target groups of this new book series. In: E. Störmer & U. Schubert (Eds.), *Sustainable Development in Europe. Concepts, Evaluation and Applications*. Cheltenham: Edward Elgar Publishing.
- HOPPE, R. (2011): *The governance of problems: puzzling, powering and participation*. Bristol: The Policy Press.
- HUMMELBRUNNER, R. (2007): *Systemic evaluation in the field of regional development*. Point Reyes: Edge Press.
- HUMMELBRUNNER, R. (2011): Systems thinking and evaluation. In: *Evaluation*, 17(4). doi: 10.1177/1356389011421935, pp. 395–403.
- KEMP, R.; ROTMANS, J. (2009): Transitioning policy: co-production of a new strategic framework for energy innovation policy in the Netherlands. In: *Policy Sciences*, 42(4). doi: <http://dx.doi.org/10.1007/s11077-009-9105-3>, pp. 303–322.
- MEADOWCROFT, J. (2007): Who is in charge here? Governance for sustainable development in a complex world. In: *Journal of Environmental Policy & Planning*, 9(3–4). doi:10.1080/15239080701631544, pp. 299–314.
- MEYER, W. (2007): Evaluation of sustainable development: a social science approach. In: E. Störmer; U. Schubert (Eds.), *Sustainable Development in Europe. In: Concepts, Evaluation and Applications*. Cheltenham: Edward Elgar Publishing, pp. 33–50.
- MIDGLEY, G. (2000): *Systemic intervention*. New York: Springer.

- MILLER, T. R. (2012): Constructing sustainability science: emerging perspectives and research trajectories. *Sustainability Science*. doi: 10.1007/s11625-012-0180-6, pp. 1–15.
- MONEVA, J. M.; ARCHEL, P.; CORREA, C. (2006): GRI and the camouflaging of corporate unsustainability. *Accounting Forum*, 30(2). doi: <http://dx.doi.org/10.1016/j.accfor.2006.02.001>, pp. 121–137.
- PATTON, M. Q. (2008): *Utilization-focused evaluation*. Thousand Oaks: Sage.
- PATTON, M. Q. (2011): *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. New York: Guilford Press.
- PATTON, M. Q. (2013): The roots of utilization-focused evaluation. In: M. C. Alkin (Ed.), *Evaluation Roots: A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: SAGE, pp. 293–303.
- POPE, J.; ANNANDALE, D.; MORRISON-SAUNDERS, A. (2004): Conceptualising sustainability assessment. *Environmental impact assessment review*, 24(6). doi: 10.1016/j.eiar.2004.03.001, pp. 595–616.
- REYNOLDS, M.; WILLIAMS, B. (2012): Systems thinking and equity-focused evaluations. In: M. Segone (Ed.), *Evaluation for equitable development results Evaluation Working Papers*. New York: UNICEF. Retrieved from http://www.capacity.org/capacity/export/sites/capacity/documents/topic-readings/Evaluation_for_equitable_results_web.pdf. Retrieved: 22.07.2013
- ROBINSON, J. (2004): Squaring the circle? Some thoughts on the idea of sustainable development. *Ecological Economics*, 48(4). doi: 10.1016/j.ecolecon.2003.10.017, pp. 369–384.
- ROTMANS, J. (2006): Tools for integrated sustainability assessment: a two-track approach. In: *Integrated Assessment*, 6(4), pp. 35–57.
- STOCKMANN, R. (2011): *A Practitioner Handbook on Evaluation*. Cheltenham: Edward Elgar Publishing.
- STOKER, G. (1998): Governance as theory: five propositions. In: *International Social Science Journal*, 155, pp. 17–27.
- STÖRMER, E.; SCHUBERT, U. (2007): Evaluation of sustainable development in Europe: context and introduction. In: E. Störmer & U. Schubert (Eds.),

- Sustainable Development in Europe. Concepts, Evaluation and Applications. Cheltenham: Edward Elgar Publishing, pp. 1–30.
- VOSS, J.-P.; BAUKNECHT, D.; KEMP, R. (2006): Reflexive governance for sustainable development. Cheltenham: Edward Elgar Publishing.
- VOSS, J.-P.; KEMP, R. (2005): Reflexive governance for sustainable development-Incorporating feedback in social problem solving. Paper presented at the ESEE Conference, June 14–17, 2005, Lisbon.
- WCED (1987): Our common future. Report on the World Conference on the Environment and Development. Oxford: Oxford University Press.
- WEAVER, P. (2013): Presentation at ICIS, Maastricht University: Integrated sustainability assessment.
- WEAVER, P.; ROTMANS, J. (2006): Integrated sustainability assessment: What is it, why do it and how? In: International journal of innovation and sustainable development, 1(4), doi: 10.1504/IJISD.2006.013732, pp. 284–303.
- WEAVER, P.; ROTMANS, J.; TURNPENNY, J.; HAXELTINE, A.; JORDAN, A. (2007): Methods and tools for integrated sustainability assessment (MATISSE): A new European project. In: C. Georges & C. C. H. Kirkpatrick (Eds.), Impact assessment and sustainable development: European practice and experience. Cheltenham: Edward Elgar Publishing, pp. 149–163.
- WEISS, C. H. (1993): Where politics and evaluation research meet. In: American Journal of Evaluation, 14(1). doi:10.1177/109821409301400119, pp. 93–106.
- WILLIAMS, B.; HUMMELBRUNNER, R. (2011): Systems concepts in action: A practitioner's toolkit. Stanford, CA: Stanford University Press.
- WILLIAMS, B.; IMAM, I. (2007): Systems concepts in evaluation. An expert anthology. Point Reyes, CA: EdgePress.

Seit von Foerster können wir soziale Systeme als Black Boxes betrachten, zu deren Funktionsweise keine klaren Wenn-Dann-Aussagen möglich erscheinen und deren Operationsweise sich für den Beobachter immer nur sequenziell und sprunghaft – und folglich nie ganzheitlich – darstellt. Jedes Bemühen, ein soziales System tiefgründig verstehen und abbilden zu wollen, kann somit schnell ein Gefühl von Orientierungslosigkeit und Überforderung auslösen – ähnlich, wie während einer Achterbahnfahrt. Für die Evaluationsdebatte resultiert aus dieser Sichtweise die Kernfrage, wie nun also im Rahmen von Evaluationen mit sozialer Komplexität umgegangen werden kann.

An diese Frage anknüpfend stellt der vorliegende Band das Feld der systemischen Therapie- und Beratungsansätze als inspirierenden Fundus vor, aus welchem sich Konzepte, Methoden und Techniken zur Gestaltung von Evaluationsvorhaben ableiten lassen. Aber welche Möglichkeiten und Grenzen offenbaren sich dabei? Lässt sich sozialer Komplexität mit Hilfe dieser Ansätze besser begegnen? Welche Rahmenbedingungen sollten dabei erfüllt sein und wie lassen sich systemische von nicht-systemischen Ansätzen unterscheiden?

