SSOAR

Open Access Repository
www.ssoar.info

Die Konstruktion des Untersuchungsgegenstandes in der Evaluation innovativer multizentrischer Programme

Haubrich, Karin

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:
Haubrich, K. (2006). Die Konstruktion des Untersuchungsgegenstandes in der Evaluation innovativer multizentrischer Programme. In K.-S. Rehberg (Hrsg.), *Soziale Ungleichheit, kulturelle Unterschiede: Verhandlungen des 32. Kongresses der Deutschen Gesellschaft für Soziologie in München. Teilbd. 1 und 2* (S. 3872-3881). Frankfurt am Main: Campus Verl. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-142469

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Die Konstruktion des Untersuchungsgegenstandes in der Evaluation innovativer multizentrischer Programme

Karin Haubrich

Evaluation zwischen Forschung und professioneller Kunst

Die Ad-hoc-Gruppe »Evaluationsforschung und Evaluationspraxis« stellt eine Frage in den Mittelpunkt des Interesses, die das grundlegende Selbstverständnis von Evaluation im Kern berührt. Zwar gilt inzwischen als weithin anerkannter Konsens, dass Evaluation ein systematisches und transparentes Vorgehen zur Bestimmung des Wertes eines Gegenstandes auf der Grundlage empirisch gewonnener Informationen darstellt (vgl. Deutsche Gesellschaft für Evaluation 2002: 13). Ergänzend wird zudem häufig als Kennzeichnen benannt, dass Evaluationen meist in Anwendungskontexten stattfinden und darauf ausgerichtet sind, praktische Maßnahmen zu überprüfen, zu verbessern oder über sie zu entscheiden (vgl. Wottawa/Thierau 2003: 14). Ob Evaluation jedoch als »anwendungsorientierte empirische Sozialforschung« (Kromrey 1995: 315) bzw. als »Teilbereich empirischer Forschung« (Bortz/Döring 2002: 101) verstanden wird, als Transdisziplin des Bewertens (Scriven 2003: 19), deren Wurzeln weiter zurückreichen als die der empirischen Sozialforschung oder als professionelle oder praktische »Kunst« (Cronbach 1982: 1; Widmer 2000: 77) hängt davon ab, welches der zentralen Kennzeichen von Evaluation – Bewertung, Anwendung von Forschungsmethoden, Praxisbezug (vgl. Haubrich/Lüders 2004) – als Ausgangspunkt einer Definition oder anspruchsvoller: einer Evaluationstheorie (vgl. Shadish u.a. 1991: 36ff.) fokussiert wird.

Ohne an dieser Stelle ausführlich diskutieren zu können, wie die »Gräben« zwischen Wissenschaft einerseits, professioneller Bewertung andererseits und schließlich praktischer Kunst oder professioneller Praxis verlaufen und mit welchen Argumenten sie verteidigt werden, tauchen die Spannungsverhältnisse zwischen den jeweils damit verbundenen Gütekriterien unweigerlich in jeder Evaluation auf und werfen methodische, methodologische und Design-Fragen auf. Solche Fragen werden in diesem Beitrag am Beispiel durchgeführter Evaluationen von Bundesmodellprogrammen im Bereich der Kinder- und Jugendhilfe diskutiert. Insbesondere wird der Frage nachgegangen, wie innovative multizentrische Modellprogramme als

komplexe Untersuchungsgegenstände theoretisch gefasst und mit Hilfe von logischen Modellen zur Herausarbeitung der impliziten Programmtheorien praktisch (re-)konstruiert werden können. Damit soll bildlich gesprochen ein Beitrag geleistet werden, die »Gräben« zwischen Evaluation als sozialwissenschaftliche Forschung und Evaluation als praktische Kunst etwas besser zu »auszuleuchten« und Wege ihrer Überbrückung zu finden.

Dafür ist es zunächst Voraussetzung, die typischen Merkmale von Bundesmodellprogrammen als Evaluationsgegenstände zu bestimmen. In einem zweiten Schritt wird der Ansatz der Cluster-Evaluation als eine praxisorientierte Antwort auf die Herausforderungen eines schwierigen Untersuchungsgegenstandes in seinen Grundzügen beschrieben und offene methodologische Fragen problematisiert. Über die Herausarbeitung der impliziten Annahmen des Evaluationsansatzes über den Untersuchungsgegenstand wird ein Vorschlag entwickelt, wie Modellprogramme als Evaluationsgegenstände (re-)konstruiert werden können. In diesem Kontext wird die Anwendung logischer Modelle als heuristische Instrumente zur Gegenstandsbeschreibung vorgestellt.

Bundesmodellprogramm als dorniges »Evaluandum«

Modellprogramme des Bundes im Bereich der Kinder- und Jugendpolitik sind exemplarische Felderprobungen im Dienste der überregionalen Anregungsfunktion des Bundes (§ 80 Abs. 1 SGB VIII; siehe auch Art. 30 GG) gegenüber der Fachpraxis und zum Zweck der Gewinnung von Erkenntnissen für die Ausgestaltungen gesetzlicher Initiativen (vgl. Dietzel/von Troschke 1988; Lüders 2003). Sie sind zeitlich befristet und dienen der Erprobung fachlicher Strategien und Arbeitsformen, die so (noch) nicht Bestandteil der Regelpraxis in diesem Politik- und Praxisfeld sind. Sie zielen auf eine Weiterentwicklung der Fachlichkeit in der Kinder- und Jugendhilfe, indem sie fundierte Erkenntnisse und Anregungen für die Implementation bewährter neuer Arbeitsformen oder die Weiterentwicklung bestehender Konzeptionen auf lokaler Ebene ermöglichen. Um die an verschiedenen Orten in Einzelprojekten eines Modellprogramms gewonnenen Erfahrungen zu bündeln und das generierte Wissen zu sichern, wird diesen Programmen eine wissenschaftliche Begleitung zur Seite gestellt.

Solche innovativen Programme zeichnen sich nun gerade dadurch aus, dass zu Beginn des Programms die zu erreichenden Ziele nur in groben Zügen beschrieben werden können und die konkreten Handlungsansätze zur Erreichung dieser Ziele erst im Programmverlauf entwickelt werden. In den meisten Fällen sind Modellprogramme an mehreren Standorten angesiedelt, das heißt es handelt sich um multi-

zentrische Programme. Typischerweise wird das Programm in unterschiedlichen Projekten gezielt unter heterogenen Bedingungen umgesetzt, damit aus den Variationen der Umsetzungsformen und den jeweiligen kontextbezogenen Lösungen gelernt werden kann und eine möglichst breite Basis für Innovationsförderung gegeben ist. Variationen der Umsetzung einer Programmidee werden bei diesem Programmtypus als erwünscht angesehen und entsprechen dem Programmzweck.

Diese Offenheit der Konzeption zu Beginn des Programms und die angestrebten Variationen in der Programmumsetzung stellen die Evaluation vor die Herausforderung, den Evaluationsgegenstand so zu beschreiben, dass das Gesamtprogramm in seinen Kerndimensionen fassbar wird. Diese Anforderung erweist sich bei genauerer Betrachtung als höchst voraussetzungsreich. Denn das Programm ist zu Beginn meist nicht viel mehr als die Ausformulierung einer allgemeinen Idee oder eines (fach-)politischen Konzeptes. Selbstverständlich liegt einem Programm eine in der Regel jedoch nicht explizit formulierte »Theorie« (die jedoch nicht mit einer sozialwissenschaftlichen Theorie gleichgesetzt werden kann) zugrunde, wie die Programmelemente miteinander verknüpft sind, um bestimmte Ziele zu erreichen. Es ist zudem davon auszugehen, dass unterschiedliche Beteiligte eines Programms jeweils eigene Interpretationen der unterstellten Wirkungszusammenhänge vornehmen und wohl kaum zwei Personen zu finden sein werden, deren implizite Theorien – ohne komplexe Aushandlungsprozesse – widerspruchsfrei übereinstimmen.

Diese Tatsache ist nicht nur darin begründet, dass je nach professionellem Hintergrund, fachlicher und fachpolitischer Orientierung die angestrebten Ergebnisse sowie die wirksamen Aktivitäten anders eingeschätzt und gewichtet werden. Vielmehr bedingt auch das Prozedere der Ausschreibung und Umsetzung von Programmideen vielfältige Interpretationsschritte: Programmideen entstehen in politischen und fachpolitischen Kontexten unter Berücksichtigung unterschiedlichster Interessenlagen der an der Programmentwicklung Beteiligten – wie Ministerium, Verbände, durchführende Organisationen oder beratende Experten. Die Programmausschreibung ist in aller Regel bewusst sehr offen formuliert, um eine große Bandbreite lokal angepasster und häufig innovativer Konzepte zu ermöglichen. In der Konsequenz sind die Antragskonzepte üblicherweise heterogen und reflektieren in erster Linie die lokale Perspektive. Die Auswahl einer bestimmten Zahl an Projekten ist der erste strategische Schritt im Sinne der Programmsteuerung, um Schwerpunkte der Programmaktivitäten zu akzentuieren. Das heißt das Auswahlverfahren und die dabei zu Grunde liegenden Auswahlkriterien konkretisieren die ursprünglich vage Programmidee. Zudem übernehmen in den geförderten Projekten oftmals Mitarbeiterinnen und Mitarbeiter die Aufgaben der Programmimplementation, die an den Antragstellungen und den anfänglichen Verhandlungen über die Programmumsetzung nicht mitgewirkt hatten. Ihre Umsetzungsideen, Interpretationen und Handlungsschritte bestimmen die faktische Konturierung des Pro-

gramms. In diesem Prozess fließen strukturell unterschiedlich gelagerte Perspektiven der Interpretation der Programmidee ein, die in den seltensten Fällen deckungsgleich sind.

Cluster-Evaluation – eine pragmatisches Evaluationskonzept

Trotz der beschriebenen Merkmale des Untersuchungsgegenstandes ist die Evaluation von Bundesmodellprogrammen gefordert, nach Abschluss des Programms gültige Aussagen über Ergebnisse, Erfolge und Misserfolge sowie ihre Ursachen auf der Programmebene zu formulieren. Dabei interessiert nicht nur die Frage, ob sich bestimmte Erfolge feststellen lassen, sondern vor allem auch warum in welchen Kontexten und in Bezug auf welche Zielgruppen bestimmte Strategien erfolgreich waren, um im Sinne einer Weiterentwicklung der Fachpraxis aus den gewonnenen Einsichten übertragbare Erkenntnisse gewinnen zu können. Zwischen sozialwissenschaftlichen Ansprüchen an methodische Designs zur Klärung von Wirkungszusammenhängen und den damit verbundenen Anforderungen an die Strukturierung des Untersuchungsgegenstandes (vgl. hierzu u.a. Kromrey 1995) einerseits und den Interessen an der Ausgestaltung des Programms und der Nutzung der Evaluationsergebnisse andererseits klafft jedoch eine unübersehbare Lücke. In der Evaluationspraxis finden sich daher unterschiedliche Evaluationsdesigns, denen der Versuch gemeinsam ist, mit begründeter Plausibilität bestimmte Ergebnisse auf bestimmte Interventionen zurückzuführen – ohne (quasi-)experimentelle Wirkungskontrollen durchführen zu können. Solche Erhebungsdesigns werden in aller Regel für die jeweilige Begleituntersuchung individuell entwickelt und nur in seltenen Fällen in ihren konzeptionellen Grundlagen und evaluationstheoretischen Orientierungen einer breiteren Öffentlichkeit zugänglich gemacht. Dies gilt mit Einschränkung auch für den Ansatz der Cluster-Evaluation, der in den USA Anfang der 1990er Jahre im Rahmen längerfristiger, auf systemische Veränderungsprozesse auf kommunaler Ebene angelegter sozialer Reformprogramme der W.K. Kellogg Foundation (vgl. Sanders 1995: 5) entwickelt wurde.

Dieser relativ junge Ansatz der Programmevaluation wurde konzipiert als Antwort auf die speziellen Anforderungen multizentrischer Programme, die sich dadurch auszeichnen, dass die einzelnen Projekte an den verschiedenen Standorten relativ autonom sind und jeweils eigene kontextbezogene Strategien entwickeln (vgl. Sanders 1997: 396). Als »Cluster« wird dabei eine Gruppe von Projekten definiert, die eine gleiche Zielsetzung verfolgen, eine gemeinsame Strategie einsetzen, dieselbe Zielgruppe haben (Fitzpatrick u.a. 2004: 476), die gleiche Konzeption zu Grunde

legen oder auch gemeinsam auf einen systemischen Veränderungsprozess in Institutionen und Fachpolitik hinwirken (W. K. Kellogg Foundation 1995: 5).

Da Cluster-Evaluation für die Lösung konkret beschreibbarer Praxisanforderungen und nicht so sehr als ein Beitrag zur Weiterentwicklung der evaluationstheoretischen Debatte konzipiert worden ist, gibt es bislang nur wenige wissenschaftliche Veröffentlichungen zu den methodologischen Grundlagen des Evaluationsdesigns. Auch wurden Cluster-Evaluationen in verschiedenen Programmevaluationen, die damit gearbeitet haben, in unterschiedlicher Weise ausgearbeitet. Ihre Dokumentationen sind kaum öffentlich zugänglich. Nach Blaine Worthen und Constance Schmitz lassen sich lediglich bestimmte Schlüsselmerkmale dieses Evaluationsansatzes benennen.

1. »It looks accross a group of projects to identify common threads and themes that, having cross-confirmation, take on greater significance;
2. It seeks not only to learn ›what happened‹ with respect to a group of projects, but ›why those things happened‹;
3. It happens in a collaborative way that allows all players – projects, the Foundation, and external evaluators – to contribute to and participate in the process so that what is learned is of value to everyone, and
4. The relationship between the projects and the external evaluators conducting the cluster evaluation is confidential. Information is reported back to the Foundation in aggregate (...).« (Worthen/Schmitz 1997: 303)

Diese Schlüsselmerkmale geben wenig Auskunft über die konkrete methodische Herangehensweise, sondern beschreiben eher einen Rahmen für die Gestaltung von Beziehungen und Rollen der beteiligten Akteure und eine Philosophie der Gewinnung, Rückvermittlung und Verwendung von Evaluationsergebnissen. Im Hinblick auf die Verfahren der Beschreibung des Gegenstandes, der Erhebung und Auswertung von Daten findet man hingegen nur als allgemein formulierte Merkmale, dass Cluster-Evaluationen holistisch und ergebnisorientiert (*outcome-oriented*) angelegt sein sollen und übertragbare oder sogar generalisierbare Erkenntnisse gewinnen sollen. Darüber hinaus gibt es einige wenige Anregungen für die Anlage des Evaluationsdesigns, die sich an den Prinzipien qualitativer Sozialforschung orientieren (vgl. Sanders 1997: 399). Die Frage, wie das zu untersuchende Programm als kohärenter Evaluationsgegenstand definiert und damit einer Analyse zugänglich gemacht werden kann, wird jedoch nicht näher erläutert und schon gar nicht methodologisch expliziert. Dafür soll in einem ersten Schritt rekonstruiert werden, was im Falle der Cluster-Evaluation unter einem Programm zu verstehen ist und in welcher Beziehung die heterogenen Projekte zur ursprünglichen Programmidee und dem Gesamtprogramm stehen.

Implizite Programmannahmen der Cluster-Evaluation

Zunächst einmal meint Programm als Terminus der Evaluationsfachsprache lediglich, dass auf der Grundlage ausgewiesener Ressourcen eine geplante Intervention umgesetzt wird, um bestimmte Ziele zu erreichen (vgl. Fitzpatrick u.a. 2004: 54). In dieser Bestimmung des Untersuchungsgegenstandes unterscheidet sich Programmevaluation von anderen Evaluationsformen wie zum Beispiel Produktevaluation, Politikevaluation oder Evaluation von Organisationen oder Personen. Jedoch umfasst diese Definition zugleich ein sehr heterogenes Spektrum von Programmen als Gegenstand von Evaluationen: Angefangen von komplexen multizentrischen und möglicherweise sogar internationalen Interventionsprogrammen mit einer Vielzahl von lokalen Projekten, über eine spezifisches Projekt bist hin zu einer einzelner Trainingseinheit.

Obwohl die oben beschriebenen Merkmale für jedes der genannten Beispiele gültig sind und auch zur Beschreibung des Gegenstandes grundsätzlich herangezogen werden können, reicht dies noch nicht aus um die Spezifika unseres »dornigen« Evaluandums fassen zu können. Der typische Charakter innovativer multizentrischer Programme besteht ja gerade darin, dass es zunächst eine allgemeine Programmidee gibt oder eine »rhetorische Programmtheorie« (Shadish u.a. 1991: 251), die dann über verschiedene Interpretationsschritte auf der lokalen Ebene in verschiedenen Projekten (tatsächliche oder reale Programmtheorie) oftmals sehr unterschiedlich umgesetzt wird. Somit würde auch eine Beschreibung der internen Programmstruktur (Shadish u.a. 1991: 37), das heißt der Komponenten wie Personalausstattung, Klienten, sachliche und finanzielle Ressourcen, Ergebnisse, Verwaltung, interne Mittelallokation und interne Organisation und der Art und Weise in der diese miteinander verknüpft sind aus Sicht unterschiedlicher Programmbeteiligter nicht einheitlich erfolgen. Nun ist die Differenz zwischen rhetorischer Programmtheorie und tatsächlicher Programmtheorie kein Phänomen, dass allein Bundesmodellprogramme betrifft, sondern genau genommen für jedes Programm in mehr oder weniger ausgeprägter Weise zutrifft. Jedoch finden wir im Ansatz der Cluster-Evaluation eine bestimmte konzeptionelle Rahmung dieses Phänomens, die sich von Programmannahmen andere Evaluationsansätze unterscheidet:

- Erstens: Wenn Projektinformationen vertraulich sind, kann nur in aggregierter Form über Erfolge und Misserfolge berichtet werden (s.o. Schlüsselmerkmal 4). Dies geht einher mit dem Fokus der Cluster-Evaluationen auf das, was aus dem Programm insgesamt gelernt werden kann und nicht auf die Frage, wie gut ein einzelnes Projekt ist.
- Zweitens: Um zu übertragbaren Ergebnisse zu gelangen, werden gemeinsame Themen der Projekte identifiziert und vergleichend betrachtet um induktive

Schlüsse über erfolgreiche oder weniger erfolgreiche Ansätze und ihre Kontextbedingungen ziehen zu können (s.o. Schlüsselmerkmal 1).

Für das Evaluationsdesign bedeutet dies, dass die Projekte nicht als Einzelprojekte, sondern als unterschiedliche Fälle der Realisierung einer Programmidee für die Untersuchung relevant sind. Die Aktivitäten der Projekte werden als Strategien oder Arbeitsformen des Programms interpretiert. Das heißt, dass bestimmte Projektaktivitäten, die keinen erkennbaren Bezug zu den relevanten Dimensionen der Programmidee haben, auch nicht Gegenstand der Analyse sind. Wenn also beispielsweise ein Projekt zur Vernetzung von Einrichtungen selbst mit Jugendlichen arbeitet, so ist dies zunächst nicht Bestandteil der Evaluation. Wenn aber deutlich wird, dass die Arbeit mit den Jugendlichen als in hohem Maße relevant für die Vernetzungsarbeit angesehen wird, weil dies als Voraussetzung für die Akzeptanz als kompetenter Partner der Jugendarbeit in der Vernetzungsstruktur gesehen wird, dann kann es zum Bestandteil der Evaluation werden. Dies kann aber nicht von der Entscheidung eines einzelnen Projektes abhängen, sondern angesichts des partizipativen Anspruchs der Cluster-Evaluation (s.o. Schlüsselmerkmal 3) nur im Zusammenwirken aller relevanter Programmbeteiligten. Wenn eine solche Arbeitsform als relevante Programmdimension angesehen wird, wird umgekehrt die ursprüngliche Idee des Gesamtprogramms modifiziert. Die – zunächst rhetorische – Programmtheorie wird also nicht nur Schritt für Schritt durch die faktischen Umsetzungsformen konkretisiert und spezifiziert, sondern auch verändert.

Das Programm lässt sich daher aus Sicht der Evaluation nur dann als kohärenter Untersuchungsgegenstand beschreiben, wenn im Zusammenwirken der unterschiedlichen Programmbeteiligten ein minimaler Konsens über die zentralen Aspekte des Programms und die Annahmen über relevante Wirkungszusammenhänge hergestellt werden kann. Das heißt es muss über gemeinsame Diskussionsprozesse definiert werden, welche Projektaktivitäten in welcher Weise welche Dimensionen des Programms erfüllen. Dieser Prozess der Definition des Untersuchungsgegenstandes muss zudem über den gesamten Programmverlauf hinweg aufrechterhalten bleiben. Schließlich sollen die Projekte – als innovative Felderprobungen – im Programmverlauf offen bleiben für Veränderungen, damit auf sich wandelnde Rahmenbedingungen reagiert sowie aus Fehlern gelernt und eine entsprechende Konzeptanpassung vorgenommen werden kann. Daraus resultiert, dass die Beschreibung des Untersuchungsgegenstandes veränderungsoffen angelegt sein muss und Evaluationen nicht durchgängig mit einem vorab festgelegten Evaluationsdesign arbeiten können, sondern vielmehr auf die entsprechenden Veränderungen im Programmverlauf durch Anpassung der Fragestellungen und Erhebungsmethoden reagieren müssen. Die *implizite Programmannahme* der Cluster-Evaluation lässt sich daher wie folgt beschreiben: Das Programm wird im Wechselwirkungsprozess

zwischen gemeinsamer Programmidee und heterogenen lokalen Umsetzungsformen durch das Handeln der Akteure vor Ort und im Kommunikationsprozess zwischen den Programmbeteiligten konstitutiert.

Ein innovatives multizentrisches Programm lässt sich also immer auf zwei Ebenen beschreiben. Einerseits auf der Ebene der Programmvision, die zunächst als rhetorische Theorie des Gesamtprogramms vorliegt und sukzessive mit der Realität abgestimmt wird. Andererseits auf der Ebene der Projekte, die im Kontext der Cluster-Evaluation nicht als ganzheitliche Fälle im Sinne einer Fallstudie sondern als Implementationen der Programmidee verstanden werden. Wenn man den Evaluationsgegenstand theoretisch in dieser Form fasst, dann stellt sich die Frage, wie dieser empirisch untersucht werden kann.

Logische Modelle als Basis des Evaluationsdesigns – die Evaluationspraxis

In diesem Zusammenhang erwiesen sich in der Praxis der Umsetzung von Cluster-Evaluationen so genannte logische Modelle (*logic models)* als hilfreiche Instrumente des Evaluationsdesigns. Sie werden eingesetzt, um die meist implizit bleibenden »Handlungstheorien« (vgl. Patton 1997: 234ff.) des Gesamtprogramms wie auch der Projekte zu explizieren. Diese Modelle basieren auf der zuvor gegebenen Definition, dass mit einem Programm auf der Grundlage ausgewiesener Ressourcen über aufeinander abgestimmte Aktivitäten bestimmte Ziele erreicht werden sollen. Sie bilden in einer Kette von »Wenn-Dann-Aussagen« ab, aufgrund welcher *Inputs* (wie finanzielle oder personelle Ressourcen), mit welchen *Aktivitäten*, welche *Outputs* (Produkte) bereitgestellt und welche *Outcomes* (Ergebnisse) und langfristigen gesellschaftlichen *Wirkungen* (*Impacts*) angestrebt werden. Sie sind in diesem Sinne vereinfachte und systematische Abbildungen der internen Programmstruktur.

Logische Modell helfen, die grundlegende Struktur eines Programms sowohl auf der Ebene des Rahmenkonzeptes als auch auf Projektebene aus Sicht der beteiligten Akteure abzubilden. Daher eignen sie sich für die Schaffung einer gemeinsamen Interpretationsbasis der Programmbeteiligten, auf deren Grundlage die Aktivitäten der Projekte unter der Programmperspektive eingeordnet, diskutiert und analysiert werden können. Logische Modelle dienen einerseits als Grundlage der Verständigung und Aushandlung zwischen den Programmbeteiligten über die Bausteine des Programms und ihre Verknüpfungen. Andererseits dienen sie als Interpretationsfolien für die Evaluation, um das Programm, wie es im Wechselwirkungsprozess zwischen vorgegebenem Programmrahmen und lokalen Umsetzungsformen durch das

Handeln der Akteure vor Ort und im Kommunikationsprozess zwischen den Programmbeteiligten konstituiert wird, rekonstruieren zu können.

Dieses Instrument ist in hohem Maße nützlich, um die Programmperspektive bei den lokalen Projekten wach zuhalten und somit dazu beizutragen, dass die Projekte sich in ihrer Heterogenität nicht bis zur *Unkenntlichkeit* des Programms auseinander entwickeln. Damit wird durch das Evaluationsinstrument unweigerlich Einfluss genommen auf die Entwicklung des Evaluandums. Mit diesen Ausführungen wird bereits deutlich, dass logische Modell vor allem als Instrumente der Evaluationspraxis entwickelt worden sind. So ist es auch nicht weiter verwunderlich, dass es zwar in der Zwischenzeit eine Reihe guter Praxisanleitungen gibt, in welchen Schritten und mit welchen Akteuren man logische Modelle entwerfen sollte. Jedoch findet sich nahezu keine eingehende Diskussion der damit verbundenen methodologischen Fragen. Hier besteht noch Bedarf einen weiteren Baustein zur Überbrückung der »Gräben« zwischen Evaluationsforschung und Evaluationspraxis zu entwickeln.

Literatur

Bortz, Jürgen/Döring, Nicola (2002), *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*, Berlin/Heidelberg/New York.

Cronbach, Lee J. (1982), *Designing Evaluations of Educational and Social Programs*, San Francisco/Washington/London.

Deutsche Gesellschaft für Evaluation (Hg.) (2002), *Standards für Evaluation*, Köln.

Dietzel, Gottfried T. W./von Troschke, Jürgen (Hg.) (1988), *Begleitforschung bei staatlichen Modellprojekten – strukturelle und methodische Probleme*, Stuttgart u.a.

Fitzpatrick , Jody L./Sanders, James R./Worthen, Blaine R. (2004), *Program Evaluation. Alternative Approaches and Practical Guidelines*, Boston u.a.

Haubrich, Karin/Lüders, Christian (2004), »Evaluation – mehr als ein Modewort?«, *Recht der Jugend und des Bildungswesens. Zeitschrift für Schule, Berufsbildung und Jugenderziehung*, Jg. 57, H. 3, S. 316–337.

Kromrey, Helmut (1995), »Evaluation. Empirische Konzepte zur Bewertung von Handlungsprogrammen und die Schwierigkeiten ihrer Realisierung«, *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie*, Jg. 15, H. 4, S. 313–336.

Lueders, C. (2003), »Jugendhilfeforschung zwischen Wissenschaft, Politik und Praxis«, *DJI Bulletin*, H. 64, S. 4–5.

Patton, M.Q. (1997), *Utilization – Focused Evaluation*, Newbury Park/London/New Dehli.

Sanders, James R. (1995), »Cluster Evaluation. A Creative Response to a Thorny Evaluation Issue«, *Newsletter of the AEA Nonprofits and Foundations TIG*, S. 5.

Sanders, James R. (1997), »Cluster Evaluation«, in: Chelimsky, Eleanor/Shadish, William (Hg.), *Evaluation for the 21st Century. A Handbook*, Thousand Oaks, S. 396–404.

Scriven, Michael (2003), »Evaluation in the New Millennium. The Transdisciplinary Vision«, in: Donaldson, Steward I./Scriven, Michael (Hg.), *Evaluating Social Programs and Problems. Visions for the New Millennium*, London u.a., S. 19–41.

Shadish, William R. Jr./Cook, Thomas D./Leviton Laura C. (1991), *Foundations of Program Evaluation. Theories of Practice,* Newbury Park/London/New Dehli.

Widmer, Thomas (2000), »Qualität der Evaluation«, in: Stockmann, Reinhard (Hg.), *Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder*, Opladen, S. 77–102.

W. K. Kellogg Foundation (1995), *Cluster Evaluation Model of Evolving Practices*, Battle Creek, MI.

Worthen, Blaine R./Schmitz, Constance C. (1997), »Conceptual Challenges Confronting Cluster Evaluation«, *Evaluation*, Bd. 3, H. 3, S. 300–319.

Wottawa, Heinrich/Thierau, Heike (2003), *Lehrbuch Evaluation*, Bern u.a.