

Rezension zu:

Giel, Susanne: Theoriebasierte Evaluation. Konzepte und methodische Umsetzungen. Münster u.a.: Waxmann, 2013. 306 Seiten, 34,90 EUR, Internationale Hochschulschriften, Band 584, ISBN 978-3-8309-2855-3

*Christian Erzberger*¹

Während vor gut 15 bis 20 Jahren der Begriff „Evaluation“ in der Wissenschaft weitgehend unbekannt war, hat er inzwischen inflationären Charakter angenommen. Jede noch so unsystematische Untersuchung wird heute mit dem Begriff der Evaluation wissenschaftlich ‚veredelt‘. Dies kann zum einen erfolgen, weil der Begriff nicht geschützt ist, und zum anderen, weil die Kriterien zur Bestimmung seines Inhaltes – z.B. die Standards der DeGEval – Gesellschaft für Evaluation e.V. – nicht bekannt sind. Darüber hinaus ‚leidet‘ die Tätigkeit der Fachkräfte, die sich beruflich mit Evaluationen beschäftigen, unter der Tatsache, von der ‚richtigen‘ universitären Wissenschaft nicht anerkannt zu sein. Dabei entsteht der Widerspruch, dass gerade Evaluationen nachhaltigere Veränderungen in Teilen der gesellschaftlichen (manchmal auch sehr kleinräumigen) Praxis auslösen als groß angelegte, DFG-geförderte Projekte der Universitäten. Die häufig anzutreffende etwas ‚herablassende‘ Haltung der Universitäten zur Evaluation hat auch damit zu tun, dass Evaluationen in der Regel als wenig theoretisch begründet angesehen werden und ihnen ein reiner Empirismus angedichtet wird, der nicht den herrschenden methodischen und methodologischen Standards entspricht.

Auch wenn diese Einleitung etwas überdimensioniert erscheint, so verweist sie doch auf eine wichtige Intention der Arbeit, nämlich die Tatsache, dass häufig Evaluationen ohne eine Rekonstruktion der zu bewertenden Programme nicht auskommen, dass es also notwendig ist, die ‚Theorie‘ eines Programms zu kennen, um überhaupt eine Evaluation entsprechend angemessen durchführen zu können. Unter diesem Aspekt sind qualitativ gute Evaluationen wahrscheinlich nie atheoretisch, da die

dem Programm unterliegenden Annahmen den Ausgangspunkt für die empirische Analyse bilden. Gleichwohl wird in der hier vorliegenden Arbeit – aus guten Gründen – der Rekonstruktion der Programmtheorie besondere Aufmerksamkeit gewidmet, was letztendlich auch Folgen für den gesamten Evaluationsprozess hat. Die Arbeit wendet sich daher fundiert gegen die Annahme, Evaluationen seien rein empiristisch und nicht theoretisch fundiert.

Die übergeordnete Fragestellung der Arbeit lautet entsprechend: „Wie können Evaluationen gegenstandsangemessen durchgeführt werden, so dass sie für Beteiligte und Betroffene einen Nutzen erzeugen, realistisch umsetzbar sind, zu genauen Ergebnissen führen und fair den Beteiligten gegenüber sind?“ (S. 11). Um dieses Ziel zu gewährleisten, so wird in der Arbeit argumentiert, müssen die den Programmen zugrunde liegenden Programmtheorien rekonstruiert und die Implikationen, die diese Rekonstruktionen auf den Evaluationsprozess haben, beachtet werden. Letzteres gilt vor allen Dingen für die Explikation der Bewertungsbasis und die methodische Umsetzung der Erhebung und Analyse. Das bedeutet, dass erst durch die Rekonstruktion eine Bewertungsbasis geschaffen wird und dass mit Blick auf den Erhebungs- und Auswertungsprozess eine methodische Offenheit vonnöten ist. Letztere wiederum muss der Programmtheorie und den Bewertungskriterien angemessen sein.

Zur Verwendung gelangt in der Arbeit der Begriff „Programmtheorie“. Das erklärt sich dadurch, dass im Mittelpunkt der Arbeit Programmevaluationen stehen, „da sie die größte Verbreitung vorzuweisen haben, weil sie umfassend funktionieren und sie letztlich implizit Produkte und Organisationsstrukturen mit erfass-

1 Gesellschaft für innovative Sozialforschung und Sozialplanung e.V. (GISS), Bremen

sen können müssen“ (S. 22). Damit wird auch erklärt, was unter Programmtheorien zu verstehen ist: die dem Programm zugrunde liegenden Annahmen sowie Hypothesen über das Programm. „Theoriebasierte Evaluation proklamiert für sich nicht nur, Auskunft darüber zu geben, ob ein Programm greift, sondern wie und weshalb dies (nicht) gelingt“ (S. 11).

Der Darstellung des Vorgehens der Rekonstruktion der Programmtheorie und der sich daraus ergebenden Folgerungen wird in der Arbeit sehr dezidiert und praxisnah nachgegangen.

Im Kapitel 2 „**Evaluation – Was ist das eigentlich?**“ werden die Aufgabenfelder von Evaluatorinnen und Evaluatoren sowie die an sie gerichteten Anforderungen mit Blick auf unterschiedliche Einsatz- und Arbeitsfelder beschrieben. Dabei werden Begriffsdefinitionen gegeben, grundsätzlich der Ablauf eines Programms und einer Evaluation dargestellt, Bewertungsdimensionen verdeutlicht, Konflikt- und Spannungsfelder von Evaluationen angesprochen und die DeGEval-Standards vorgestellt. Darüber hinaus wird in ein Beispiel einer eigenen Evaluation im Bereich der Einführung einer internetbasierten Lernunterstützung für das Lehrangebot „Einführung in die Methoden der empirischen Sozialforschung“ eingeführt. Dieses Beispiel wird in allen folgenden Kapiteln zur Demonstration der praktischen Umsetzung der beschriebenen Verfahren genutzt.

Das Kapitel 3 dient zur Darstellung der „**Klassische[n] Konzepte zur Durchführung von Evaluationen**“. Dies geschieht, um die Differenz zu den theoriebasierten Konzepten zu verdeutlichen. Als relevant werden hier Konzepte angesehen, die „grundlegende Schulen und Paradigmen von Evaluationen repräsentieren“ (S. 12). „Demzufolge liegt die Herausforderung darin, die zentralen Konzepte herauszufiltern, um die Bandbreite der verschiedenen Perspektiven abzubilden und dabei gleichzeitig die wichtigsten Einflüsse – vor allem für die deutschsprachige Evaluation – zu rekonstruieren“ (S. 55). Näher beschrieben werden dabei das zielbasierte Konzept von Ralph Tyler, die experimentellen und quasiexperimentellen Designs von Campbell und Stanley – als vermeintlicher Goldstandard von Evaluationen –, die nutzungsfokussierte Evaluation von Patton und die konstruktivistische Evaluation von Lincoln und Guba. Es werden jeweils die Grundzüge der Konzepte beschrieben und die Rolle der Evaluatorinnen und Evaluatoren darin beleuchtet. Den Abschluss bildet eine kritische Würdigung des jeweiligen Konzeptes, wobei hier auf den An-

satz der „zielfreien Evaluation“ von Scriven zurückgegriffen wird.

Im Kapitel 4 „**Konzepte theoriebasierter Evaluationen**“ werden die Alternativen zu den klassischen Konzepten beschrieben. „Mit theoriebasierten Konzepten findet Evaluation einen alternativen Zugang zu ihrem Gegenstand: Weder werden Ziele als Ausgangspunkt gewählt noch startet die Evaluation mit einem spezifischen methodischen Design (...). Vielmehr nähert sich die Evaluation ihrem Gegenstand über die im Programm zugrunde liegenden Annahmen und Hypothesen“ (S. 137). Entsprechend werden hier zunächst die Schwächen der klassischen Konzepte abgehandelt, um damit die Grundlage für die Beschreibung der theoriebasierten Evaluation zu schaffen. Auch wird hier näher auf den Begriff der Theorie eingegangen und dieser präziser als „Programmbezogene Theorie“ bezeichnet. Dabei wird davon ausgegangen, dass Programmtheorien entlang des Evaluationsgegenstandes konzipiert werden. Dabei kann auf externe Theorien (z.B. Lerntheorien) zurückgegriffen werden oder auf interne Theorien über eine vermutete Wirkung eines Programms. Nicht selten können mehrere Theorien für ein Programm extrahiert werden. Da es für diese Extrahierung nicht nur ein Konzept gibt, werden unterschiedliche Autorinnen und Autoren herangezogen und deren Vorstellungen zur Erlangung von Programmtheorien dargestellt. Das Kapitel endet mit der Beantwortung der Frage, welche Funktion die Programmtheorien für die Evaluation und deren Nutzen für das Programm haben.

Nachdem die Konzepte vorgestellt wurden, stehen im Kapitel 5 „**Methodische Umsetzung theoriebasierter Evaluationen**“ die Probleme und Herausforderungen bei der praktischen Umsetzung der Konzepte im Mittelpunkt. Dabei werden verschiedene methodologische Vorstellungen zur Überprüfung herangezogen. In den Blick geraten hier die naturalistische Forschung und der Symbolische Interaktionismus, die Aktionsforschung, die Dokumentarische Methode, die Grounded Theory und standardisiert verfahrenende Forschungsstrategien, die dem Kritischen Rationalismus verpflichtet sind (z.B. Kausalmodelle, Pattern Matching). „Diese verschiedenen Forschungsrichtungen werden jeweils auf ihre wissenschaftstheoretischen Grundannahmen und auf ihr methodisches Vorgehen hin analysiert, um fundierte Auskunft über deren Nutzungsmöglichkeiten für theoriebasierte Evaluationen treffen zu können“ (S. 13). Als Ergebnis wird in diesem Kapitel festgehalten, dass in der Regel kein methodisches Vorgehen alleine in der Lage

ist, den Anforderungen einer programmtheoretisch fundierten Evaluation gerecht zu werden.

Entsprechend dieser Erkenntnis geht es im Kapitel 6 „**Integration von Methoden zur Umsetzung theoriebasierter Evaluationen**“ um die Kombinationen unterschiedlicher Methoden, Daten und Ergebnisse. Die Entscheidungen für bestimmte methodische Vorgehensweisen sind demnach an die zuvor ermittelte Programmtheorie zurückgebunden und müssen sich nach ihren Erfordernissen richten. Das bedeutet, dass mit der theoriebasierten Evaluation eine methodische/methodologische Offenheit zwangsweise einhergeht. „Eine puristische Haltung, wonach sich aufgrund entgegenstehender wissenschaftstheoretischer Paradigmen die Kombination unterschiedlicher methodischer Zugänge verbietet, kann sich die angewandte Sozialforschung und erst recht die Evaluation nicht leisten“ (S. 242). In diesem Kapitel werden daher Möglichkeiten des Methodenmix, der Triangulation und der Methodenintegration vorgestellt und diskutiert.

Eine Zusammenfassung der Arbeit und ein Ausblick finden sich im Kapitel 7 „**Potentielle theoriebasierter Evaluationskonzepte und zu lösende Herausforderungen**“. Das Besondere an dieser Arbeit wird hier nochmals verdeutlicht: „Nun könnte man einwenden, dass selbstverständlich jeder Art von Evaluation eine Programmtheorie zu Grunde liegt. So scheinen zumindest auf den ersten Blick Konzepte theoriebasierter Evaluation nichts wirklich Neues zu bieten. Ein zweiter Blick verdeutlicht jedoch, dass die Programmtheorie, die den entscheidenden Referenzrahmen für die theoriebasierte Evaluation liefert, genau und detailliert expliziert werden muss. Damit nicht genug: Theoriebasierte Evaluation schärft die Aufmerksamkeit für die (Daten-)Grundlage von Programmtheorien“ (S. 274). Die Explikationen der internen und externen programmtheoretischen Annahmen sind nur über den Einbezug von Stakeholdern möglich und liefern dann Grundlagen für die methodischen Zugänge und für die Bestimmung der Bewertungskategorien. Eine Programmtheorie

wird nach diesem Modell nicht logisch abgeleitet, sondern empirisch, datenbasiert und methodisch kontrolliert entwickelt.

Auch wenn aufgrund des Titels der Arbeit sich zunächst möglicherweise Personen angesprochen fühlen, die bereits über Erfahrungen in der Evaluation verfügen, so ist die Arbeit unbedingt auch für ‚Einsteiger‘ zu empfehlen. Dies nicht nur, weil im ersten Teil auf die Grundbegriffe von Evaluationen eingegangen wird, sondern auch, weil alle theoretischen und methodologischen/methodischen Darstellungen immer an einem durchlaufenden Beispiel sehr praktisch erläutert werden. Insgesamt ergibt sich dadurch eine interessante Verquickung von theoretischer Beschreibung und deren praktischer Anwendung. Weiterhin außerordentlich positiv zu bewerten ist die mit der Explikation einer Programmtheorie verbundene methodologische/methodische Offenheit von Evaluationen. Hier wird deutlich, dass durch den Gegenstand die Methoden definiert werden und nicht umgekehrt. Das bedeutet zum einen, dass die Evaluatoreninnen und Evaluatoren über gute methodische Kenntnisse verfügen müssen, um diesem Umstand gerecht werden zu können – es bedeutet aber auch, keine Methode von vornherein als absolut zu setzen und unterschiedliche wissenschaftstheoretische Paradigmen als unüberwindbar anzusehen. Diese in der Arbeit verfolgte Argumentation kann nicht positiv genug bewertet werden, feiert doch die Beharrung auf geradezu antagonistischen Differenzen zwischen unterschiedlichen Methodologien/Methoden im Bereich der Universitäten fröhliche Urständ. Aber noch ein weiterer Vorteil der theoriebasierten Evaluation wird hier herausgearbeitet. Sie ermöglicht die Explikation von Bewertungskategorien. Damit wird die Ziellinie, ob eine Evaluation ihren Annahmen gerecht geworden ist, ex ante definiert. Dies ist ein großer Vorteil gegenüber Evaluationen, die die Wirkungen erst ex post ermitteln – also starten, ohne zu wissen, wo der ‚Erfolgspunkt‘ liegt. Insgesamt eine überaus kenntnisreiche Arbeit, die zudem noch sehr gut lesbar ist.

Rezension zu:

Stockmann, Reinhard/Silvestrini, Stefan (Hg.): *Metaevaluierung Berufsbildung. Ziele, Wirkungen und Erfolgsfaktoren der deutschen Berufsbildungszusammenarbeit*. Münster u.a.: Waxmann, 2013. 208 Seiten, 29,90 EUR, Reihe Sozialwissenschaftliche Evaluationsforschung, Band 12, ISBN 978-3-8309-2795-2

*Philipp Gonon*¹

Die vorliegende Studie, herausgegeben und im Wesentlichen verfasst von Reinhard Stockmann und Stefan Silvestrini, fällt insofern aus dem Rahmen üblicher Evaluationen, als sie auf einer Evaluation von Evaluationen basiert. Wie die Autoren richtigerweise bemerken, sind Metaevaluationen selten im Einsatz und kaum Gegenstand theoretischer und methodologischer Reflexion. Gerade im bildungspolitischen Diskurs und im Rahmen von evidenzbasierter Steuerung erhalten solche Untersuchungen allerdings vermehrt Aufmerksamkeit. So wird bereits in der Einleitung auf den Zusammenhang zur „Evidenz“-Basierung als Forderung, die inzwischen auch die Entwicklungszusammenarbeit erreicht hat, hingewiesen. Immerhin gab und gibt es in diesem Bereich Querschnittsanalysen, die verschiedene Einzelevaluierungen quer vergleichend auswerten. Die beiden Herausgeber gebrauchen beide Begriffe, wie auch zusätzlich den Term Synthesebericht, und heben hierbei die inhärent gemeinsame zentrale Zielsetzung hervor, kumuliertes Wissen zur Verfügung zu stellen. Querschnittsanalyse und Metaevaluierung, so die Hoffnung, sollen also zusätzliche Informationen aus Einzelevaluierungen generieren, nicht nur für einen Sektor oder eine Region, sondern generell auch zu einer Verbesserung hinsichtlich Evaluierungsmethoden beitragen. Der Mehrwert der Metaanalyse wird insbesondere darin gesehen, durch eine vergleichende Auswertung anhand selbstgewählter Fragestellungen zusätzlich neues Wissen zu erzeugen.

Methodisch orientiert sich diese Untersuchung – im Unterschied zu vielen Untersuchungen nach diesem Ansatz textanalytisch anhand von Berichtsdokumenten – an der Grounded Theory; dies obwohl ein Kategoriensystem weit-

gehend vorgegeben war, das dann jedoch induktiv, das heißt also anhand des Datenmaterials modifiziert und ergänzt wurde. Die methodischen Überlegungen werden alles in allem sehr knapp gehalten, ohne hierbei die Grenzen eines solchen Ansatzes oder die besonderen Vorzüge gegenüber Alternativen näher zu beleuchten.

Es geht also einmal um das Identifizieren von Mustern und übergreifenden Merkmalen. Dafür ‚eignet‘ sich die Entwicklungszusammenarbeit im besonderen Maße, ist sie doch seit Jahrzehnten mit einer Fragestellung konfrontiert, die an die ‚Substanz‘ geht: Bewirkt sie das, was sie zu bewirken vorgibt, überwiegen unerwünschte Effekte oder sind gar keine nennenswerten Folgen zu erkennen? Augenfällig gilt dieser generelle Vorbehalt auch für die Berufsbildungszusammenarbeit mit Entwicklungsländern: Gerade „duale Projekte“, die im Besonderen neben Schule auch die betriebliche Seite der Ausbildung im Blick haben, sowie deren Einführung sind in ihrer Bilanz rückwirkend nicht sehr erfolgreich gewesen, dennoch wird weiter gefördert; sei es aufgrund der Tatsache, dass man um Effekte nichts weiß, wissen will oder fehlende Erfolge als tolerierbar betrachtet.

Das Buch ist in insgesamt drei Teile gegliedert. Nach einer allgemeinen Darstellung von Begrifflichkeiten und Funktionen der Evaluation, bzw. der Querschnittsauswertungen und Metaevaluationen im Rahmen der GIZ (Gesellschaft für Internationale Zusammenarbeit), folgt ein ausführlicher zweiter Teil, der die Einzelauswertungen von 12 Förderprojekten im Querschnitt auswertet, ehe eine Metaevaluierung folgt, die abschließend mit Folgerungen und Empfehlungen abgerundet wird. Ein dritter prägnanter und knapp gehaltener Teil ist

1 Lehrstuhl für Berufsbildung, Institut für Erziehungswissenschaft, Universität Zürich

dann mit „Management Response“ überschrieben, gibt also die Reaktion der Auftraggebenden bzw. Nachfragenden wieder. Dem im Vorwort erwähnten Anspruch, die Einbettung der Studie transparent und auch für Außenstehende gut nachvollziehbar zu machen, wird damit Genüge getan. Wie festgehalten, werden im ersten Teil neben den Begrifflichkeiten in einem zweiten und dritten Unterkapitel – verfasst von Stefanie Krapp und Antje Schmidt – darüber hinaus die Funktionen und Fragestellungen der Evaluationen im Zusammenhang mit der GIZ dargestellt.

Zunächst entfaltet Stefan Silvestrini eine Querschnittsanalyse der 12 Projekte (Sierra Leone, Bulgarien, Indien, Indonesien, Kasachstan, Usbekistan, Brasilien, Burkina Faso, China, Ghana, Mazedonien und die Philippinen). Nicht ganz klar wird hierbei, wie es gerade zu dieser Auswahl kam, außer dass diese Evaluationen – eine als Interimsevaluation, andere als Schlussequalierungen oder Ex-post-Evaluierungen erstellt – alle im Jahre 2010 erfolgten. Sie entsprächen gemäß Fußnote etwa einer Stichprobengröße von 15 % der durchgeführten Evaluationen (S. 29). Entlang von sieben Kriterien, die u.a. die Ziele und Handlungsfelder, aber auch die Berücksichtigung und Bewertung von DAC-Kriterien, wie Impact und Effizienz, sowie die auf Berufsbildung beziehbaren Millenniumsziele (MDG), Fachlichkeit und nachhaltige Entwicklung einschließen, werden diese Projekte dargestellt und gemäß den Bewertungen der Primärevaluatorinnen und -evaluatore „gescreent“. Tatsächlich wird dadurch vergleichend ein Profil der unterschiedlich ambitionierten und ausgestatteten Projekte sichtbar. In übersichtlichen Tabellen werden dadurch Kooperationspartner, Zielgruppen, Erfolgs- und Misserfolgskriterien auf der Ebene der Umsetzung, der Institutionen und des Systems verdeutlicht. Im vergleichenden Querschnitt wird auch herausgearbeitet, dass unterschiedliche Semantiken und Verständnisse zu Begriffen wie „Breitenwirkung“ und Nachhaltigkeit bei den Begutachterinnen und Begutachtern viele Fragen und auch Lücken in der Beurteilung offen lassen, zumal sich oft auch eine übergreifende strategische Orientierung im internationalen Geberkontext nicht erschließen lasse (S. 99). Zur Wirkungsattribution sind dann – kaum überraschend – keine näheren Aussagen eruierbar, was Silvestrini in der Zusammenfassung zur Querschnittsanalyse zu einer Kritik der Berichterstattungen veranlasst.

Der titelgebende anschließende Beitrag von Stockmann nimmt sich u.a. dann gerade vor, diese zuvor konstatierten Lücken und Män-

gel näher zu beleuchten. Es werden zunächst die konzeptionelle Entwicklung der Berufsbildung und im Besonderen dann auch der Einfluss der Millenniumsziele auf die Berufsbildungsförderung nachgezeichnet. Die anschließenden Kapitel fokussieren die Wirksamkeit und Nachhaltigkeit, um weiter auf Charakteristika und Erfolgsfaktoren einzugehen. Auch durch den Einbezug zusätzlicher und zeitlich zurückliegender Evaluationsstudien gelingt es ihm, Wandel und eine jüngere Geschichte der beruflichen Entwicklungszusammenarbeit zu rekonstruieren, was insofern naheliegend ist, als er selbst auch Autor dieser Studien in den 1990er Jahren war. Als interessanter Befund sei festgehalten, dass vonseiten der Evaluatorinnen und Evaluatoren Effizienz und Impact durchgängig besser beurteilt werden als die Nachhaltigkeit, was auf gewisse Widersprüchlichkeiten in den Bewertungspraktiken hinweist. Der Bezug zu den Förderzielen sei in den Projekten häufig recht vage und es ist von begrifflichen Umetikettierungen bzw. Labeling-Effekten die Rede, um nicht konkreter auf Vorgaben eingehen zu müssen. Kritisch wird insgesamt konstatiert, dass die Bewertungshandlungen der Evaluatorinnen und Evaluatoren nicht konsistent seien, da unterschiedlichste und uneinheitliche Methoden und Bewertungsprozeduren im Spiel waren (S. 158). Der Autor plädiert im Schlussteil seiner Metaanalyse aufgrund der 59 identifizierten Einflussgrößen für ein „Balanced Score Card Modell“, um die Schlüsselgrößen für nachhaltige Entwicklungsarbeit zu verorten. Dieses aus der Managementtheorie übernommene Modell, welches die Vision und Strategie im Zusammenhang mit bzw. als Wechselspiel von Ownership, flexibler Steuerung, Personal und Systemkompatibilität sieht, ergibt sich nicht zwingend aus den vorgängigen Ausführungen, erscheint wohl aber praktikabel. Dieses Modell wurde daher wohl auch nicht in den die gesamte Untersuchung zusammenfassenden Schlussfolgerungen weiter aufgenommen und ausgeführt und es wird ebenso wenig in der Management Response vonseiten der GIZ aufgegriffen.

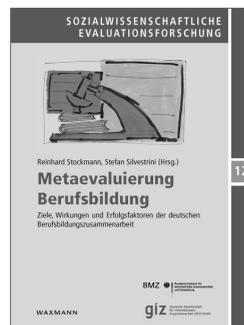
Schlussfolgernd konstatieren die Herausgeber in ihrer synthetischen Darstellung eine Überforderung hinsichtlich der gesetzten sehr allgemein gehaltenen entwicklungspolitischen Zielsetzungen und verweisen aufgrund der Analyseresultate auf Defizite bei der Umsetzung der Projekte. Als ein Ergebnis wird festgehalten, dass Institutionenförderprojekte mit einem begrenzten Anspruchsprofil eine höhere Nachhaltigkeit erzielen würden.

Was bedeuten diese Aussagen angesichts der anfänglich formulierten Ansprüche an die Metaevaluation? Ähnlich wie bei gehobener kulinarischer Zubereitung ist auch die Metaanalyse angewiesen auf gute Produkte, bzw. Grundlagen, die erst ein gelungenes Menü ergeben. Auch in dieser vorliegenden Evaluation wird diese Ausgangslage, die auf unterschiedliche Qualitätsprodukte aufbauen muss, an einer Stelle problematisiert, wenn es etwa um die Einschätzung der Nachhaltigkeit geht. Das Schwanken der Beurteilungskriterien der Primärevaluatorinnen und -evaluatorenen und die Heterogenität der Projekte selbst erschweren eine konsistente Analyse und das Anliegen, weiterführende Fragestellungen

auf die Projekte selbst anzuwenden und darauf bauend auch neue Evaluationskriterien zu formulieren. Insgesamt ist die Datenlage eher begrenzt, um weitreichende Folgerungen zu ziehen. Dennoch sind das Vorgehen und auch der Anspruch dieser Querschnittsanalyse und Metaevaluation sehr fruchtbar und eröffnen weitere Fragestellungen und Evaluationsperspektiven.

Insofern ist diese Studie bedeutsam und bemerkenswert, als in den letzten Jahren die Politik der Entwicklungszusammenarbeit die Berufsbildung wiederentdeckt hat. Bestehende Evaluationen, die kritisch das vorhandene Wissen beleuchten, sind daher auf jeden Fall erwünscht und hilfreich.

Im Auftrag des Bundesministeriums für wirtschaftliche Zusammenarbeit und Entwicklung (BMZ) führt die Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) seit einigen Jahren Querschnittsauswertungen ihrer Einzelevaluierungen durch. Neu ist dies in Kombination mit einer Metaevaluation. Erstmals werden hier die Vorgehensweise, Erkenntnisse, Empfehlungen und die Verwertung einer solchen Studie vorgestellt.



Reinhard Stockmann,
Stefan Silvestrini (Hrsg.)

Metaevaluierung Berufsbildung

Ziele, Wirkungen und
Erfolgsfaktoren der
deutschen Berufsbildungs-
zusammenarbeit

2013, 208 Seiten, br., 29,90 €
ISBN 978-3-8309-2795-2
E-Book: 26,99 €



WAXMANN

Rezension zu:

Dziekian, Katrin/Riedel, Veronique/Müller, Stephanie/Abraham, Michael/Kettner, Stefanie/Daubitz, Stephan: *Evaluation Matters: A Practitioners' Guide to Sound Evaluation for Urban Mobility Measures*. Münster u.a.: Waxmann, 2013. 180 Seiten, 18,90 EUR, ISBN 978-3-8309-2881-2

Jörn Gründler,¹ Christoph Emanuel Müller¹

Eine einfache Möglichkeit des Kompetenzerwerbs in der Evaluation besteht in der Konsultierung von Fachliteratur. Insbesondere im Einsteigersegment, dem auch das vorliegende Buch zuzuordnen ist, existiert eine Vielzahl von Werken, deren schiere Masse kaum mehr zu überblicken ist. Es stellt sich daher die Frage, ob ein weiteres Einführungsbuch überhaupt benötigt wird. Diese Frage wird durch das Buch „*Evaluation Matters*“ grundsätzlich mit ‚ja‘ beantwortet! Dies liegt nicht nur daran, dass der Text einen Themenbereich aufgreift, der bislang kaum Gegenstand evaluationsbezogener Einführungsliteratur war. Vor allem handelt es sich beim Buch um eine verständlich geschriebene – und dennoch keinesfalls anspruchslose – Lektüre für all diejenigen, die mit der Entscheidung, Planung, Umsetzung, Analyse, Bewertung oder Weiterentwicklung urbaner Transportinterventionen zu tun haben. Primäre Zielgruppe sind grundsätzlich alle Personen, die sich für die Evaluation im Themenbereich der urbanen Mobilität interessieren.

Das zentrale Anliegen des Buches besteht darin, anhand theoretischer Konzepte und empirischer Beispiele einen grundlegenden Überblick über das Instrument der Evaluation (mit Fokus auf den Themenbereich der urbanen Mobilität) zu geben. Das Buch beginnt mit einem **Einführungskapitel**, in welchem zunächst das Konzept der Evaluation definiert und von anderen Konzepten wie denen des Monitorings/Audits abgegrenzt wird. Weiter werden potenzielle Zielsetzungen von Evaluationen vorgestellt und mögliche Evaluationszeitpunkte im Projektimplementierungsprozess thematisiert. Die Einführung schließt mit der Beschreibung potenzieller Stakeholder von Evaluationen und beantwortet

die aufgeworfene Frage „Why do we need this book?“ mit den folgenden Worten: „[...] there is still a lack in evaluation literature that is specifically tailored to the urban mobility-related context.“

Im **zweiten Kapitel** wenden sich die Autor(inn)en der wichtigen Thematik der Wirkungsevaluation zu. Hierbei definieren sie zunächst eingängig, was es mit Wirkungsevaluationen auf sich hat, wozu diese umgesetzt und wie sie organisiert werden können. Es wird darauf hingewiesen, weshalb die saubere Konzeptualisierung der Ziele einer Intervention für den Erfolg ihrer Wirkungsevaluation unerlässlich ist. Darüber hinaus erläutern die Verfasserinnen und Verfasser die Wichtigkeit der Entwicklung von Programmtheorien, welche sie als „cause-and-effect-chains“ bezeichnen und anhand intuitiv nachvollziehbarer Beispiele veranschaulichen. Die im Kapitel bislang erarbeiteten Erkenntnisse münden in Empfehlungen zur Ableitung von Indikatoren, die die Outcome-Variablen abbilden, auf denen potenzielle Wirkungen gemessen werden sollen. Bevor in einem zweiten Teil des zweiten Kapitels auf eher praktische und umsetzungsbezogene Aspekte der Wirkungsevaluation eingegangen wird, wird ein kurzer Überblick über die für Wirkungsevaluationen so wichtigen Untersuchungsdesigns gegeben. In verständlicher Art und Weise werden die gängigen Elemente von Untersuchungsanordnungen (Kontrollgruppen und Vorhermessungen) beschrieben, anhand derer in der Evaluationspraxis der hypothetische kontrafaktische Zustand geschätzt wird. Die unterschiedlichen Formen der Nutzung dieser Elemente (beispielsweise in Form von randomisierten Experimenten, Quasi-Experimenten oder Pretest-Posttest-Designs) werden

¹ Centrum für Evaluation (CEval), Saarbrücken

in einer zusammenfassenden Heuristik präsentiert.

Im Nachgang zur Präsentation der Untersuchungsanordnungen werden eher praktische Komponenten von Wirkungsevaluationen thematisiert. Beginnend mit der Datenerhebung wird in einem ersten Schritt auf unterschiedliche Datenquellen (Primär- und Sekundärdaten) sowie auf Aspekte der Datentriangulation eingegangen. Nachfolgend werden in zielgruppengerechter Sprache die Datenerhebungsstrategien Tiefeninterview, Fragebogenumfrage, Fokusgruppeninterview und Beobachtung beschrieben. Trotz der einsteigerfreundlichen Formulierungen könnte es für manche Leser(innen) an dieser Stelle etwas knifflig werden, die vorgestellten Methoden in den Kontext der Wirkungsevaluation einzuordnen, da die vorgestellten qualitativen Datenerhebungsstrategien nicht unbedingt mit den im Vorfeld beschriebenen Designs in Verbindung zu bringen sind. Für rigorose Wirkungsevaluationen werden standardisierte Indikatoren benötigt, welche i.d.R. nicht durch die Anwendung qualitativer Erhebungsverfahren generiert werden. Ergänzend zur Nebeneinanderstellung der Erhebungsstrategien wäre daher der Hinweis angebracht gewesen, dass im Rahmen der kausalanalytischen Wirkungsbewertung auf Standardisierung ausgerichtete Erhebungsverfahren, zur Identifikation von Erfolgsfaktoren aber auch qualitative, offenere Verfahren eingesetzt werden sollten/könnten.

Im Folgenden werden kurz unterschiedliche Erhebungsformen wie Face-to-Face-Interviews, Telefonumfragen, postalische Befragungen und Online-Befragungen vorgestellt und deren Vor- und Nachteile umrissen. Es folgt ein kurzer Überblick über die Problematik der Stichprobenziehung. Hier wird deutlich, mit welcher schwieriger Aufgabe die Verfasser(innen) des Buchs konfrontiert sind: die Reduktion komplexer und abstrakter Konzepte auf ein Niveau, welches auch von Einsteiger(inne)n zu erfassen ist. Während die zentralen Determinanten statistischer Konfidenz anschaulich beschrieben werden, hätte die gesamte Thematik der Stichprobenziehung konsistenter in den Gesamtkontext der vorangegangenen Inhalte eingeordnet werden können, da die hier vorgestellte und auf generalisierende Analysen ausgerichtete Art der Stichprobenziehung in aller Regel nicht im Rahmen qualitativer oder Fokusgruppeninterviews zur Anwendung kommt. Ein Vorwurf ist den Autor(innen) allerdings nicht zu machen, da eine umfassendere Beschreibung sowohl Rahmen als auch angestrebtes Niveau des Buchs sicherlich auch gesprengt hätte.

Das Kapitel zur Wirkungsevaluation endet mit einem Abschnitt zur Datenauswertung. Positiv zu würdigen ist hierbei die Strategie der Verfasser(innen), den Leser(inne)n bereits zu Beginn des Abschnitts zu vermitteln, dass durch das Buch aufgrund der Komplexität quantitativer und qualitativer Auswertungsverfahren lediglich eine eher oberflächliche Einführung gegeben werden kann. Es wird zu Recht darauf hingewiesen, dass in der Praxis ggf. Expert(inn)en hinzugezogen werden sollten. Im Kontext quantitativer Auswertungsverfahren konzentrieren sich die Autor(innen) zunächst auf Aspekte der Datenaufbereitung und -sicherung. Anschließend werden die grundlegendsten Verfahren der deskriptiven Statistik (Häufigkeiten, Lagemaße, Streuungsmaße) vorgestellt. Es folgt eine nachvollziehbare Erläuterung des Konzepts der Inferenzstatistik, wobei positiv zu würdigen ist, dass die zentrale Voraussetzung der Zufallsauswahl explizit thematisiert wird. Obwohl die vorgestellten Verfahren lediglich Grundlagen darstellen und in der Regel nicht annähernd für statistische Analysen im Rahmen rigoroser Wirkungsevaluationen ausreichend sind, so ist die Strategie, die Leser(innen) auf die entsprechende Einführungsliteratur zu verweisen, aufgrund der Komplexität vieler quantitativer Auswertungsverfahren sicherlich angebracht. Zum Abschluss des zweiten Kapitels wird noch eine kurze Einführung in die qualitative Datenanalyse gegeben. Dabei wird zunächst auf die Transkription und die Datenaufbereitung qualitativer Daten eingegangen. Im Anschluss daran wird kurz beschrieben, was es mit der Codierung qualitativer Daten auf sich hat und wie entwickelte Codes bzw. Codesysteme verglichen und generalisiert werden können. Schließlich widmet sich ein kurzer Schlussabschnitt der Diskussion der aus quantitativen und qualitativen Datenanalysen gewonnenen Resultate. Im Vordergrund steht hierbei insbesondere die Frage der Interpretation der Ergebnisse vor dem Hintergrund der anfänglich erläuterten Programmziele und der in Form von „cause-and-effect-chains“ postulierten Ursache-Wirkungs-Beziehungen.

Im **dritten Kapitel** des Buchs rückt nun eine andere Art von Evaluation in den Fokus des Interesses: die Prozessevaluation. Hierbei werden zunächst die Ziele und angestrebten Ergebnisse von Prozessevaluationen beschrieben. Dabei wird unter anderem näher erläutert, dass sich Prozessevaluationen vorwiegend auf Implementierungs- und Entwicklungsprozesse von Interventionen beziehen und dass die Einbeziehung unterschiedlicher Stakeholderperspektiven von großer Wichtigkeit ist. Fortfolgend be-

schreiben die Autor(inn)en die zentrale Aufgabe einer Prozessevaluation: die Identifikation von fördernden Faktoren („drivers“) und hemmenden Faktoren („barriers“) bei der Implementierung einer Maßnahme. Für beide Arten von Einflussfaktoren werden zahlreiche Beispiele aus dem Bereich der urbanen Mobilität genannt. In einem kurzen Abschnitt wird hier anschließend erläutert, wie im Evaluationsprozess gewonnene Informationen über hemmende und fördernde Faktoren sowohl in die Steuerung des Projekts als auch in die Evaluation einfließen können. Ein Hauptabschnitt des dritten Kapitels befasst sich in der Folge mit Datenerhebungsmethoden, welche in der Prozessevaluation zum Einsatz kommen. Hierbei werden zwei methodische Ansätze, standardisierte Fragebögen und „Learning Histories“, im Detail erläutert. In einem abschließenden Abschnitt stellen die Autor(inn)en die bereits erläuterten Datenerhebungsverfahren den beiden Verfahren des Fokusgruppeninterviews und des Leitfadeninterviews überblicksartig gegenüber und beschreiben deren Vor- und Nachteile in der Evaluationspraxis. Den Schwerpunkt legen die Autor(inn)en in diesem Kapitel vor allem auf die Darstellung der Methode der „Learning Histories“. Die Planung, Durchführung und Nachbereitung von „Learning History Workshops“ wird ausführlich erläutert und Literaturangaben zur Vertiefung der Erkenntnisse zur Verfügung gestellt. Hier wäre hilfreich gewesen, wenn die Autor(inn)en für die nur überblicksartig dargestellten Methoden ebenfalls Angaben über weiterführende Literatur gemacht hätten. Den Autor(inn)en gelingt es insgesamt, erfolgreich zu vermitteln, welche Wichtigkeit die Identifikation von hemmenden und fördernden Faktoren für den Erfolg schon in der Implementationsphase eines Programms besitzt und wie die generierten Informationen in die Projektsteuerung bzw. Projektkonzeption einfließen können.

Dem chronologischen Ablauf eines realen Evaluationsprojekts folgend stellen die Autor(inn)en im nachfolgenden **Kapitel vier** dar, wie Evaluationsergebnisse erfolgreich kommuniziert und nutzbar gemacht werden. Entlang von Fragen erläutern die Autor(inn)en in einem ersten Abschnitt, warum der erfolgreichen Präsentation und Dissemination von Evaluationsergebnissen eine große Bedeutung für deren spätere Nutzung zukommt. Anschließend werden die Vor- und Nachteile verschiedener Diagrammtypen diskutiert und grafisch illustriert. Hierbei konzentrieren sich die Autor(inn)en nahezu ausschließlich auf die Darstellung von quantitativen Daten. In der zweiten Hälfte des vierten Kapi-

tels gehen die Autor(inn)en auf die Nutzung der Ergebnisse von Evaluation aus Modellprojekten ein. Da Evaluationsergebnisse aus Modellprojekten im Sektor der urbanen Mobilität in der Regel in die Entscheidung über die Ausweitung in der gleichen Region („upscaling“) oder den Transfer in eine andere Region („transfer“) eines Projekts einfließen, besprechen die Autor(inn)en diese Verfahren im Detail. Eingehend wird hierbei erläutert, in welchen Situationen die Anwendung beider Verfahren angebracht ist und welche ökonomischen, institutionellen und geographische Faktoren ausschlaggebend für die Bestimmung der Erfolgsaussichten von „upscaling“- und „transfer“-Maßnahmen sind. Hilfreich ist hierbei besonders die praxisnahe Darstellung von konkreten Entscheidungsproblemen und Lösungsvorschlägen. Abschließend wird schrittweise an einem kurzen Beispiel dargestellt, welche Informationen bei der Entscheidung für oder gegen den „transfer“ einer Maßnahme zwischen zwei Städten benötigt werden und wie das notwendige Entscheidungswissen generiert werden kann.

Die Hauptzielgruppe des Buches „Evaluation Matters“ sind nicht nur angehende Evaluator(inn)en, sondern auch angehende Auftraggeber(innen) von Evaluationen, besonders jene, die eine Evaluation von städtischen Transportsystemen in Auftrag geben. In **Kapitel fünf** setzen sich die Autor(inn)en daher konstruktiv mit Argumenten auseinander, die gegen die Beauftragung einer Evaluation sprechen, um so nicht nur angehende Evaluator(inn)en, sondern auch Auftraggeber(innen) von der Effektivität und Nachhaltigkeit von Evaluationen argumentativ zu überzeugen. Diese Erläuterungen werden im Anschluss erweitert durch eine kurze Darstellung von Risikofaktoren, welche sich hemmend auf den Erfolg einer Evaluation auswirken können, wie beispielsweise unvorhergesehene Ereignisse, Zeitprobleme oder die Nichtverwertbarkeit erhobener Daten. Vor dem Hintergrund der identifizierten Risiken werden abschließend überblicksartig bewährte Verfahren vorgestellt, welche die Effektivität und Qualität eines Evaluationsprojekts gewährleisten. Insgesamt wählen die Autor(inn)en in diesem Kapitel eine sehr knappe Darstellungsweise, weshalb das Kapitel den Leser(inne)n allenfalls eine sehr selektive und vereinfachende Einführung in die Thematik bietet. Für Leser(innen), die an einer Vertiefung interessiert sind, geben die Autor(inn)en jedoch weitere Literaturempfehlungen auf Einführungsniveau an.

Im anschließenden **sechsten Kapitel** beschreiben die Autor(inn)en vor dem Hinter-

grund eigener Evaluationserfahrungen, dass valide Informationen, die im Rahmen von sorgsam durchgeführten Evaluation gewonnen werden, für Programmentscheidungen von Programmmanager(inne)n und Politiker(inne)n einen hohen Nutzen besitzen. In diesem Zusammenhang gelingt es den Autor(inn)en, überzeugend darzulegen, dass die Effektivität und Effizienz von „urban mobility“-Programmen nachhaltig gewährleistet werden kann, wenn Evaluationskomponenten bereits während des Planungsprozesses in das Programm integriert werden. In dem knapp zweiseitigen Kapitel zielen die Autor(inn)en abschließend darauf ab, die Skepsis von Programmmanager(inne)n und Auftraggeber(inne)n gegenüber Evaluationen zu reduzieren.

Das **siebte Kapitel** des Buches verbindet die in den vorherigen Kapiteln dargestellten Evaluationskonzepte mit Praxisbeispielen aus drei von den Autor(inn)en durchgeführten Evaluationen im Bereich urbaner Transportsysteme. Entlang von Textauszügen aus den Evaluationsberichten der Evaluationsprojekte („Utrecht Road Safety Label“, „Utrecht Cargohopper“, „Tallin Knittin Bus“) wird sehr anschaulich aufgezeigt, wie die in den vorherigen Kapiteln dargestellten Komponenten einer Evaluation zeitlich und inhaltlich miteinander verbunden sind. Da eines der wichtigsten Endprodukte einer Evaluation der Abschlussbericht ist, erhält der Leser/die Leserin durch die Orientierung an der Berichtsstruktur einen realistischen und praxisnahen Einblick in die Generierung und Aufbereitung von Evaluationsergebnissen. Die Evaluationsberichte sind in der Art aufgebaut, dass in einem ersten Schritt dargestellt wird, in welche Problembeschreibungen und Kontextbedingungen die Programme eingebettet sind und welche Programmziele aufgrund dessen definiert wurden. Darauf aufbauend beschreiben die Autor(inn)en die Programmkonzeption und -durchführung des jeweiligen Programms und zeigen zudem vereinzelt auf, wie Evaluationskomponenten in die Konzeption eingearbeitet sind. Im Anschluss wird unter Einsatz von hilfreichen Schaubildern der in der Evaluation stattfindende Forschungsprozess dargestellt: von der Entwicklung eines Wirkungsmodells über die Entwicklung von Indikatoren und Datenerhebungsinstrumenten und der anschließenden Datenerhebung bis hin zur Datenanalyse und Generierung von Empfehlungen. Besonders detailliert gehen die Autor(inn)en hierbei auf die Interpretation der Ergebnisse der Wirkungs- und Prozessevaluation ein. Ergänzt wer-

den die Evaluationsberichte im Anhang durch Datenerhebungsinstrumente (ein standardisierter Fragebogen sowie Formulare für die Durchführung von Prozessevaluationen), die in den dargelegten Evaluationen zur Anwendung kamen. Mit der überblicksartigen Einbettung des gesamten Evaluationsprozesses in die Struktur eines Evaluationsberichts gelingt es den Autor(inn)en sehr überzeugend, die praktische Relevanz von Evaluationsprozessen zu verdeutlichen. Auf diese Weise wird es angehenden Evaluatoren oder Auftraggeber(inne)n von Evaluationen ermöglicht, ein „abgerundetes Gefühl“ für die Evaluationspraxis zu entwickeln.

Insgesamt betrachtet bietet das Buch „Evaluation Matters“ den Leserinnen und Lesern einen einsteigerfreundlichen Überblick über die Evaluation im Bereich der urbanen Mobilität. Insbesondere die verständliche Sprache, die von den Autor(inn)en verwendet wird sowie die verständliche Darstellung teilweise komplexer Sachverhalte unterstützt Praktiker(innen) dabei, die angebotenen Inhalte ohne das vertiefte Studium von Grundlagenliteratur zu verarbeiten. Auch die verwendeten Erläuterungscomics tragen zu einem besseren Verständnis bei und lockern die Leseatmosphäre auf. Zudem ist deutlich zu erkennen, dass die Autor(inn)en über umfangreiche Kenntnisse und Erfahrungen im Bereich der urbanen Mobilität verfügen, was nicht zuletzt an den gut gewählten Beispielen zu erkennen ist.

An der einen oder anderen Stelle wäre es ggf. sinnvoll gewesen, themenspezifische weiterführende Literatur direkt an der Stelle anzugeben, an der bestimmte Sachverhalte oder Verfahren genannt wurden. Darüber hinaus ist anzumerken, dass die Autor(inn)en nicht immer in der vielleicht notwendigen Schärfe deutlich gemacht haben, dass es sich bei Evaluationen, insbesondere bei Wirkungsevaluationen, um äußerst komplexe Unterfangen handelt, die nicht unbedingt so einfach umzusetzen sind, wie an der einen oder anderen Stelle suggeriert wird.

Zusammenfassend kann aber konstatiert werden, dass es sich beim vorliegenden Buch um einen gelungenen Spagat zwischen Praxisorientierung/-anwendbarkeit und wissenschaftlicher Perspektive handelt. Für Personen, die im Bereich der urbanen Mobilität tätig sind und sich auch für die Evaluierung von Interventionen in diesem Themenfeld interessieren, bislang aber über wenig Erfahrungen mit der Evaluationsthematik verfügen, stellt das Buch eine äußerst empfehlenswerte Lektüre dar.

Rezension zu:

Knierim, Andrea/Bausch, Stefanie/Gottschick, Manuel (Hg): Partizipation und Klimawandel. Ansprüche, Konzepte und Umsetzung. München: oekom, 2013. 294 Seiten, 34,95 EUR, ISBN-13: 978-3-86581-454-8

*Wolfgang Meyer*¹

Dies ist kein Buch über Evaluation. Gleichwohl behandelt es ein für Evaluationen wichtiges Thema: die Partizipation. Während in den Evaluationszeitschriften dieses Thema eher unter ideologischen Gesichtspunkten zwischen den Polen Evaluation als Instrument zur Beteiligung (emanzipatorischer Anspruch) und Beteiligung als Instrument zur Evaluation (forschungspraktische Notwendigkeit) oder als rein technische Fingerübung bestimmter Partizipationsverfahren abgehandelt wird, bietet dieses Buch eine systematische und praxisrelevante Auseinandersetzung mit dem Problem der Beteiligung von Stakeholdern. Viele Evaluator(innen) scheinen den Willen und die Bereitschaft zur Partizipation als grundsätzlich gegeben vorauszusetzen und übersehen, dass genau dieses in der Praxis das größte Problem darstellt: Nicht nur die Beteiligungsmöglichkeiten und -chancen, sondern auch die Beteiligungsbereitschaft und die Einsicht in die Notwendigkeit, seine Interessen zu einem bestimmten Zeitpunkt zu vertreten, sind ungleich zwischen und in den einzelnen Stakeholdergruppen verteilt. Gerade für Evaluationen, die den Anspruch einer fairen Beteiligung aller Betroffenen verfolgen, ist eine methodische Absicherung der Umsetzung dieses Anspruchs von hervorragender Bedeutung und damit eine Auseinandersetzung mit den eventuell auftretenden Schwierigkeiten.

Der vorliegende Reader stellt eine systematische Darstellung der Partizipationsprobleme im Klimaschutz dar und ist aufgrund der gelungenen Aufarbeitung der Erfahrungen in diesem Anwendungsbereich auch für Leser(innen) interessant, die sich inhaltlich nicht mit Klimafragen auseinandersetzen. Grundlage ist ein am Zentrum für Agrarlandforschung (ZALF) durchge-

führter Workshop im Mai 2012. Für Auftraggebende kann dieser Band als Anregung zur Gestaltung entsprechender Diskussionen im Programmverbund dienen, denn die Beiträge dieses Buchs entstanden alle in demselben Programmkontext („KLIMZUG – Klimawandel in Regionen zukunftsfähig gestalten“ des Bundesministeriums für Bildung und Forschung – bmbf). So konnten die unterschiedlichen Erfahrungen der Projektnehmer bereits frühzeitig untereinander ausgetauscht werden. Durch die Publikation dieser Erfahrungen soll nun eine weitere Verbreitung gefördert und dadurch eine Verbesserung der Beteiligung bei Klimaschutzvorhaben erreicht werden.

Der Reader umfasst sechzehn Beiträge, die drei Abschnitten zugeordnet worden sind. Diese Gliederung leitet sich ab aus einem emanzipatorischen Verständnis von Partizipation „als ein Prozess der gleichwertigen und aktiven Beteiligung unterschiedlicher Akteure[...], der alle Phasen eines Entwicklungsprozesses oder Programmes umfasst, von der Situationsanalyse über die Planung bis hin zur Bewertung von Maßnahmen“ (S. 11). Daraus ergibt sich zunächst ein normativer Anspruch an menschliches Handeln im Kontext von Programmen, der durch die Anwendung von partizipativen Verfahren realisiert werden soll. Mit diesen normativen Aspekten beschäftigen sich die fünf Beiträge des ersten Abschnitts.

Heike Walk stellt in ihrem überaus lesenswerten Beitrag die Frage nach den verschiedenen gesellschaftlichen Funktionen von Partizipation und den daraus abzuleitenden politischen Grundströmungen bei der Realisierung von Beteiligungsverfahren. Sie unterscheidet dabei die Funktion der Interessenintegration für effiziente-

1 Centrum für Evaluation (CEval), Saarbrücken

res Handeln (Effizienzperspektive), die Verwirklichung individueller Teilhabe zur gemeinsamen Gestaltung von Gesellschaft (emanzipatorische Perspektive) sowie das Streben nach politischer Demokratisierung durch institutionalisierte Bürgerbeteiligungen (demokratische Perspektive). Aus diesen unterschiedlichen Ansprüchen an Partizipation leiten sich in der Praxis nicht nur Interessenskonflikte auf einer Metaebene ab, sondern auch die Anwendung unterschiedlicher Beteiligungsverfahren und Instrumente, die ihrerseits wiederum nur bestimmten Funktionen gerecht werden können. Dementsprechend gehört es „zu den Herausforderungen gelungener partizipativer Bürgerbeteiligung, für jeden Anwendungsfall eine angemessene Methodenform zu finden, die alle Beteiligten zufrieden stellt und eine zielführende, effiziente und faire Lösung für das vorliegende Problem liefert“ (S. 34). Dies lässt sich uneingeschränkt auch auf die Durchführung partizipativer Evaluationen übertragen.

Die beiden folgenden Beiträge von *Stefanie Baasch* und *Cordula Krupp* beschäftigen sich mit dem normativen Gehalt der gegenwärtigen Klimapolitik sowie den daraus abzuleitenden Fallstricken für Beteiligungsverfahren. Stefanie Baasch diskutiert hierbei den Umgang mit dem politischen Ziel der „Gerechtigkeit“ als Basis für die partizipative Einbeziehung der Beteiligten, während Cordula Krupp sich mit konkreten Problemen von Partizipation im Rahmen einer „Bürgerwerkstatt“ beschäftigt. In beiden Fällen wird warnend hervorgehoben, dass selbst gut gemeinte Beteiligungsverfahren aufgrund der unterschiedlichen Machtpositionen – sei es durch politische Legitimation oder erworbene Expertise – zu einer ungleichen Lastenverteilung, dem Ausschluss bestimmter (Sub-)Gruppen oder einer Fehlleitung des Partizipationsprozesses führen können.

Ein Beispiel für grundsätzlich in Beteiligungsverfahren auftretende Konflikte diskutieren *Stefan Rötzel* und *Kollegen* im nächsten Beitrag. Hier geht es um die insbesondere in komplexen Verfahren notwendige (einseitige) Wissensvermittlung, die z.B. in der Klimaschutzpolitik zur Aufklärung der Beteiligten über bestimmte Sachfragen und notwendige Verhaltensweisen oder Entscheidungen unabdingbar ist, und auf der anderen Seite die gemeinsamen und gleichberechtigten Lernprozesse aller Beteiligten im Beteiligungsprozess, die durch ein solches Wissensgefälle (und die damit häufig verbundene und von einigen Beteiligten wahrgenommene „Lehrer-Schüler“-Relation) konterkariert werden. Hieraus resultieren Konflikte („Belehrung

erzeugt Widerstand“) und Behinderungen eines für den Erfolg von Beteiligungsverfahren notwendigen iterativen Kommunikationsstils. Am Ende kann der Rückzug oder die Demotivierung bestimmter Teilnehmergruppen stehen.

Konfliktlinien dieser Art werden quasi durch den ordnungspolitischen Rahmen, in dem Beteiligungsverfahren agieren müssen, vorbestimmt und geprägt. *Simone Haffner* stellt die gegenwärtigen verwaltungsrechtlichen Gegebenheiten für eine Öffentlichkeitsbeteiligung auf den Prüfstand und kommt zu dem Schluss, dass die anhaltende Kritik an den vorhandenen Beteiligungsverfahren zeigt, „dass es bisher nicht gelungen ist, die Werthaltung und Interessen der Bevölkerung in das Verfahren zu integrieren“ (S. 93). So muss die Letztentscheidungskompetenz, bedingt durch die im Grundgesetz vorhandene Regelung, bei der Verwaltung verbleiben und die in öffentlichen Beteiligungsverfahren erarbeiteten Lösungen sind für diese Entscheidungen keineswegs bindend. Der hieraus resultierende Konflikt zwischen formaler Verfügungsgewalt und informeller Entscheidungsprädisposition beeinflusst nicht nur Bürgerbeteiligungsverfahren, sondern auch formative Evaluationsprozesse häufig sehr negativ.

Der zweite Abschnitt des Readers thematisiert die Frage nach den Akteuren, die in den Partizipationsprozess eingebunden werden sollen. Dabei steht hier bei den aufgeführten vier Beiträgen weniger die Stakeholderanalyse als Verfahren denn die Problematik der wahrgenommenen „Betroffenheit“ und die sich aus den Beteiligungsverfahren ergebenden Interaktionen im Fokus der Betrachtungen. Bereits im Beitrag von *Simone Haffner* wird das grundsätzliche Problem einer repräsentativen Abbildung der Interessen in Beteiligungsverfahren angesprochen, welches nun in dem Beitrag von *Anne Bundschuh* und *Andrea Knierim* an einem praktischen Beispiel konkretisiert wird. Der Beteiligungsprozess im Rahmen des Innovationsnetzwerks Klimaanpassung Brandenburg Berlin (INKA BB) führt zu einer Selbstselektion der teilnehmenden landwirtschaftlichen Praxispartner, deren soziostrukturelle Daten sich deutlich vom Durchschnitt der Brandenburger Landwirte unterscheiden. In einem freiwilligen Prozess ist eine solche, im Zeitverlauf zunehmende Differenzierung zwischen Beteiligung und Betroffenheit letztlich nicht zu verhindern und prägt die praktische Verwertbarkeit der erarbeiteten Resultate. Die Autorinnen betonen deshalb die Notwendigkeit einer „kritischen sozialwissenschaftlichen Begleitforschung, die die strukturellen und gesellschaftlichen Bedingungen sol-

cher Beteiligungsprozesse transparent macht“ (S. 118).

Die drei weiteren Beiträge dieses Abschnitts beschäftigen sich mit konkreten Auswirkungen der verwendeten Konzeptionen zur Einbeziehung von Akteuren in die Beteiligungsprozesse. *Sybille Bauriedl und Kolleginnen* zeigen die mit dem Begriff der „Betroffenheit“ verbundenen Auswahlprobleme auf. Insbesondere beklagen sie den Mangel vergleichender Analysen zu partizipativen Prozessen, „die qualifizierte Aussagen zu Betroffenheit und zur Identifikation von betroffenen Akteuren zulassen“ (S. 128). Hierzu könnte eine grundlagenorientierte Evaluationsforschung angesichts der vielfältigen Anwendungsbeispiele einen wertvollen Beitrag leisten. Der Beitrag von *Manuel Gottschick und Juliane Ette* hebt die räumliche Dimension der Beteiligung und die sich hieraus in den Interaktionen der Akteure wiederzufindenden Verwerfungen hervor. Die Netzwerkbildung und -entwicklung innerhalb einer Region kann sowohl im zeitlichen Verlauf als auch an den Schnittstellen zwischen verschiedenen Netzwerken zu Problemen führen, die im Rahmen der Steuerung solcher Prozesse zu beachten und durch ein geeignetes Netzwerkmonitoring zu überwachen sind. Die notwendigen Bedingungen gelingender Beteiligung stehen im Zentrum des Aufsatzes von *Gérard Hutter und Jörg Bohnefeld*. Am Beispiel des Verbundprojektes „Regionales Klimaanpassungsprogramm Modellregion Dresden“ (REGKLAM) zeigen die Autoren vier Spannungsfelder solcher Programme zur Etablierung von Governance-Netzwerken auf und diskutieren einige der beobachteten Auswirkungen auf die Beteiligungsprozesse.

Schließlich beschäftigen sich fünf Beiträge im dritten und letzten Abschnitt mit den gewonnenen Erfahrungen bei der Umsetzung verschiedener Beteiligungsmethoden und -instrumente in den unterschiedlichen Kontexten dieses Programms. *Sonja Siart und Andrea Knierim* analysieren in ihrem Beitrag die in den verschiedenen Teilprojekten des INKA BB-Programms eingesetzten Partizipationsverfahren in theoretischer und methodischer Hinsicht. Durch die Auswertung der in allen Teilprojekten durchgeführten SWOT-Analysen stellt dies in gewissem Sinn eine Metaevaluation (allerdings auf bescheidener Datenbasis) dar. Zu den auch für formative Evaluationen interessanten Befunden gehören die geschilderten Probleme bei der Vermittlung sozialwissenschaftlicher Methoden im Rahmen der Aktionsforschung, weil „die Einsicht in die Nützlichkeit solcher Investitionen [...] meist erst im Projektverlauf“ (S. 190) entsteht. Auch der

nächste Beitrag von *Patrick Scherhauser und Kollegen* zur Partizipation in regionalen integrierten Vulnerabilitätsassessments gehört zu den aus Evaluationssicht spannendsten Beiträgen dieses Readers. Vulnerabilitätsabschätzungen sollen den Akteuren helfen, die komplexen Zusammenhänge des Klimawandels zu verstehen und geeignete Maßnahmen zu implementieren. Die Autoren analysieren anhand eines theoriebasierten Kriterienkatalogs die Steuerung der Beteiligung in 14 ausgewählten Teilprojekten. Trotz der hohen Bedeutung, die der Partizipation zugerechnet wird, zeigen die Ergebnisse eine „unstrukturierte und improvisierte oder aber zumindest eine nur ungenügend dokumentierte Vorgehensweise“ (S. 211). Der Einsatz von Partizipationsverfahren erfolgt unsystematisch und nicht auf der Basis sorgfältigen Monitorings und Evaluation der Wirkungen und ist auch hinsichtlich der genutzten Verfahren nicht sonderlich vielfältig.

Die letzten drei Beiträge befassen sich schließlich mit sehr speziellen Partizipationsverfahren. *Wiebke Pohler und Christian Zottl* beschäftigen sich mit Online-Partizipations-tools und deren Einsatzmöglichkeiten. Ein wichtiger Befund ist dabei, dass die optimale Nutzung solcher Verfahren nur in Verbindung mit persönlichen Workshops möglich ist. Mit den Schwierigkeiten in Workshops speziell bei der Verwendung von Szenarien setzen sich *Thomas Zimmermann und Kollegen* anhand von drei Fallbeispielen auseinander. Aufgrund der nur begrenzten Aussagekraft ihrer eigenen Studie regen die Autoren „eine umfassendere Wirkungsanalyse bzw. Evaluation von Szenario-Workshops als partizipativem Verfahren der Klimaanpassung“ (S. 256) an. Der abschließende Beitrag schließlich fasst die Ergebnisse einer Diskussionsrunde zur Rolle der Wissenschaftler(innen) in Partizipationsprozessen zusammen. Diese von *Andrea Knierim und Kolleginnen* erstellte Dokumentation kommt zu dem Schluss, dass eine neutrale wissenschaftliche Position in einem solchen interaktiven Prozess nicht zu halten ist. Es lassen sich vier Formen von Begleitforschung mit verschiedenen Anforderungen und Vorgehensweisen unterscheiden, die in der Praxis zudem häufig miteinander vermischt sind: Grundlagenforschung, kooperative Forschung, Aktionsforschung und Evaluationsforschung. Hieraus ergeben sich multiple Rollenanforderungen an die Wissenschaftler(innen) im Partizipationsprozess, die von diesen zu bewältigen sind. Leider bleibt dieser Beitrag sehr oberflächlich und arbeitet die Konsequenzen nur ungenügend aus.

Der Band schließt mit einer Synthese der Beiträge durch den Herausgeber und die Herausgeberinnen, die drei Fragen nach Zweck und Ziel der Partizipationsverfahren, nach den beteiligten Gruppen und der Begründung ihrer Beteiligung sowie nach den Formen der Beteiligungsverfahren zu beantworten sucht. Bezüglich der ersten Frage kommen der Autor und die Autorinnen zu dem Befund, dass die Effizienzperspektive eindeutig überwiegt und die eingesetzten Partizipationsverfahren deutlich hinter den theoretisch-konzeptionell angelegten Möglichkeiten zurückbleibt. Die Auswahl der zu beteiligenden Personen und Personengruppen erfolgt oft zufällig und manchmal willkürlich: „Nur in wenigen Fällen wird beschrieben und reflektiert, mit welchen Konsequenzen sich bottom-up bzw. situationale Beteiligung von Praxispartnern und Stakeholdern auseinandersetzen muss“ (S. 274). Es mangelt somit häufig an der sorgfältigen Planung und Vorbereitung von Partizipationsprozessen. Die zentrale Herausforderung der Ge-

staltung der Beteiligung wird nur ungenügend angegangen, wobei die Probleme weniger auf der methodischen denn auf der normativ-konzeptionellen Ebene und in der fehlenden theoretischen Anbindung zu sehen sind. „Die Beiträge des Buchs zeigen, dass ein erheblicher Entwicklungsbedarf bei der Ausgestaltung von Partizipationsprozessen besteht“ (S. 277).

Diese Schlussfolgerung gilt sicher nicht nur für die Klimapolitik, die nun dank diesem Reader über einen sehr gut geschriebenen, praxisrelevanten und wissenschaftlich fundierten Überblick verfügt. Es wäre wünschenswert, wenn auch in anderen Politikfeldern solch kompetente Grundlagenanalysen vorgelegt würden. Selbst in den führenden Evaluationszeitschriften mangelt es an einer vergleichbaren Analyse zu den Problemen und Wirkungen von Partizipationsprozessen und -verfahren. Dementsprechend ist zu hoffen, dass dieses recht preiswerte Buch (34,95 EUR) entsprechende Anstöße geben kann.