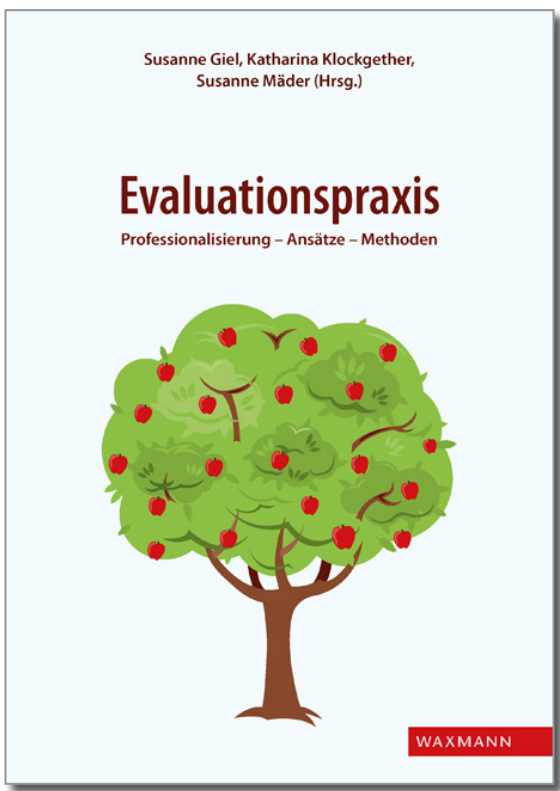


LESEPROBE

Susanne Giel,
Katharina Klockgether,
Susanne Mäder
(Hrsg.)

Evaluationspraxis
Professionalisierung –
Ansätze – Methoden

2015, 300 Seiten, br., 34,90 €,
ISBN 978-3-8309-3345-8
E-Book: 30,99 €,
ISBN 978-3-8309-8345-3



© Waxmann Verlag GmbH, 2015
Alle Rechte vorbehalten. Nachdruck, auch auszugsweise,
verboten. Kein Teil dieses Werkes darf ohne schriftliche
Genehmigung des Verlages in irgendeiner Form repro-
duziert oder unter Verwendung elektronischer Systeme
verarbeitet, vervielfältigt oder verbreitet werden.



WAXMANN

Steinfurter Str. 555
48159 Münster
Fon 02 51 – 2 65 04-0
Fax 02 51 – 2 65 04-26
info@waxmann.com
www.waxmann.com

Bestellung
Fax: 0251 26504-26
Tel.: 0251 26504-0

Internet: www.waxmann.com/buch3345
E-Mail: order@waxmann.com

Susanne Giel, Katharina Klockgether,
Susanne Mäder (Hrsg.)

Evaluationspraxis

Professionalisierung –
Ansätze –
Methoden



Waxmann 2015
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

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

Print-ISBN 978-3-8309-3345-8

E-Book-ISBN 978-3-8309-8345-3

© Waxmann Verlag GmbH, Münster 2015
Steinfurter Straße 555, 48159 Münster

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Inna Ponomareva, Jena

Titelbild: © Fictionalhead, Fotolia.de

Satz: Sven Solterbeck, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.
Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des
Verlages in irgendeiner Form reproduziert oder unter Verwendung
elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Evaluationspraxis reflektieren – Einleitung und Überblick	7
<i>Susanne Giel/Katharina Klockgether/Susanne Mäder</i>	

Professionalisierungsinstrumente

Beitrag zu einer konsistenten Evaluationsterminologie – das Glossar von Univation	19
<i>Melanie Niestroj/Wolfgang Beywl/Berthold Schobert</i>	

Checklisten zur Planung und Steuerung von Evaluationen	29
<i>Katharina Klockgether</i>	

Informationssicherheit in der Evaluation	41
<i>Marc Jelitto</i>	

Interkulturelle Kompetenz – ein weiterer Baustein für das professionelle Handeln in Evaluationen	53
<i>Schahrzad Farrokhzad</i>	

Schwarm-Evaluation als Steuerungsansatz	71
<i>Dörte Schott und Wolfgang Beywl</i>	

Evaluationsansätze

Der Programmbaum als Evaluationsinstrument	87
<i>Samera Bartsch/Wolfgang Beywl/Melanie Niestroj</i>	

Wirkungen auf der Spur mit Programmtheorien	111
<i>Susanne Giel</i>	

Developmental Evaluation nach Michael Patton Begriffsbestimmung und Reflexion der praktischen Anwendung	131
<i>Maria Gutknecht-Gmeiner</i>	

Evaluationsmethoden

Nutzen einer wirkungsorientierten Zielsystematik Bindeglied zwischen Evaluation und Programmsteuerung	157
<i>Stefan Schmidt</i>	
Der Umgang mit Monitoring in der Evaluation	175
<i>Katharina Klockgether</i>	
Interviews mit Jugendlichen Herausforderungen in der Planung und Umsetzung	197
<i>Matthias Sperling</i>	
Mit Gruppendiskussionen kollektive Orientierungen rekonstruieren	217
<i>Ute B. Schröder</i>	
Die Anwendung der Grounded Theory in Evaluationen	235
<i>Marianne Lück-Filsinger</i>	
Evaluation und das Web	253
<i>Elitsa Uzunova</i>	
Bewertungsverfahren in der Evaluation Vorgehensweisen und Praxisbeispiele	277
<i>Susanne Mäder</i>	
Autorinnen und Autoren	297



Die in diesem Band verwendeten Evaluationsfachbegriffe sind im Glossar der wirkungsorientierten Evaluation von Univation definiert. Die Begriffe sind verlinkt. Mit einem Klick gelangen Sie direkt ins Wiki: www.eval-wiki.org

Evaluationspraxis reflektieren – Einleitung und Überblick

Susanne Giel/Katharina Klockgether/Susanne Mäder

Anliegen des Buches

Evaluation, als wissenschaftliche Dienstleistung, die insbesondere öffentlich verantwortete bzw. finanzierte Politiken, Programme, Projekte oder Maßnahmen systematisch und datengestützt beschreibt und bewertet, steht kontinuierlich vor der Aufgabe, angemessene und überzeugende Vorgehensweisen einzusetzen. In der Evaluation, wie in jeder angewandten Wissenschaft, geht es nicht darum, theoretische Kenntnisse in Bezug auf den Evaluationsgegenstand zu erweitern, „sondern wissenschaftliche Verfahren und Erkenntnisse einzubringen, um sie für den zu evaluierenden Gegenstand nutzbar zu machen“ (Kromrey 2001, S. 113). Evaluationen müssen den Nutzungsabsichten, den Bedingungen und Stakeholdern der Programme gerecht werden. Dies drückt sich u. a. in den Standards für Evaluation (DeGEval 2008a), insbesondere in den Standardgruppen Nützlichkeit, Durchführbarkeit und Fairness, aus. Da Evaluationen eine Grundlage für Entscheidungen in und über Programme sind, kommt es wesentlich darauf an, nachvollziehbare und glaubwürdige Ergebnisse zu liefern. Dementsprechend sind Evaluatorinnen und Evaluatoren einem systematischen, transparenten und begründeten Vorgehen verpflichtet, so wie es in den Genauigkeitsstandards niedergelegt ist.

Dass professionelle Evaluationen auf feldspezifischen Kenntnissen, konkreten Erfahrungen und ihrer Reflektion basieren, formuliert auch der Arbeitskreis ‚Aus- und Weiterbildung in der Evaluation‘ der DeGEval in seinen Empfehlungen.

„Evaluation zeichnet sich dadurch aus, dass sie spezifische Kompetenzen voraussetzt, diese aber mehr als in anderen sozialwissenschaftlichen Feldern erst durch ihre spezifische Zusammenführung und Interpretation in der Praxis zu hinreichenden Evaluationskompetenzen werden.“ (DeGEval 2008b, S. 27).

Dementsprechend kann die Kompetenzentwicklung für Evaluatorinnen und Evaluatoren als ein kontinuierlicher Prozess aufgefasst werden, bei dem entsprechend der Anforderungen von Auftraggebenden und Erfordernissen in Programmen Lösungen entwickelt, erprobt und reflektiert werden.

Eine typische Herausforderung, der wir in unserer Evaluationspraxis häufig begegnen, bieten Evaluationsgegenstände, die als Antwort auf aktuelle Entwicklungen Neues erproben, wie modellhafte Ansätze zur Prävention in der Gesundheitspolitik oder der Jugendarbeit, innovative Ausbildungsformen oder Arbeitsmarktkonzepte. Diese Programme zeichnen sich durch ihren experimentierenden Charakter aus, der sich in Zielverschiebungen oder in im Prozess entwickelten und fortlaufend angepassten Interventionen zeigen kann. Ebenso treffen wir auf Programme, die neue

Wege in der Umsetzung beschreiten, indem sie in Netzwerken realisiert werden. Hier entstehen Programmresultate nicht nur durch die Aktivitäten Einzelner, vielmehr wird erwartet, dass gerade im Zusammenwirken verschiedener Netzwerkakteure sogenannte Synergieeffekte entstehen, die durch die Evaluation adäquat erfasst werden sollen. Einer standardisierten Vorgehensweise sind gerade im Umgang mit solchen und anderen Neuerungen deutliche Grenzen gesetzt. Vielmehr müssen in weit höherem Maße kontinuierliche Klärungsprozesse mit den Auftraggebenden, Beteiligten oder auch Zielgruppen erfolgen, um die gesamte Evaluation auf die sich stetig verändernde Programmwirklichkeit zuzuschneiden. Das Design der Evaluation muss dafür offen sein und die Methoden müssen flexibel auf die Anforderungen abgestimmt werden können.

Nicht nur auf innovative Programme, sondern auch auf allgemeine gesellschaftliche Entwicklungen müssen Evaluationen angemessen reagieren. Beispielhaft seien folgende Tendenzen genannt: Seit einiger Zeit wird eine interkulturelle Öffnung in gesellschaftlichen Institutionen und Organisationen vorangetrieben, die ihren Widerhall in unterschiedlichsten Programmen und Projekten findet. Auch Evaluatoredinnen und Evaluatoren sind herausgefordert, Vielfalt und Diversität in ihre Praxis zu integrieren. Eine nicht mehr ganz so neue Entwicklung, die jedoch gerade den Umgang mit Informationen und Wissen kontinuierlich und tiefgreifend verändert, ist die zunehmende Digitalisierung und Internetbasierung in allen gesellschaftlichen Bereichen, damit auch in der Evaluationspraxis. Es ergeben sich hiermit völlig neue Möglichkeiten der Informationsbeschaffung, der Datenverarbeitung und Ergebnispräsentation, jedoch auch neue Herausforderungen, wie z. B. ein erweiterter Datenschutz.

An diesen Ausführungen wird deutlich, dass zur Gestaltung einer angemessenen Evaluationspraxis vielfältige Entscheidungen zu treffen sind. Die konkrete Evaluationspraxis – und damit Entscheidungen über Evaluationsansätze, Evaluationsmethoden und die Steuerung der Evaluation – basiert in jedem einzelnen Fall auf einem innovativen und kreativen Prozess, in dem immer eine situativ passende Antwort auf aktuelle Herausforderungen gefunden werden muss.

Zentrales Anliegen dieses Buches ist, die kreativen und situativen Entscheidungen, die sich in der Praxis stellen, sichtbar zu machen, um so einerseits für die Vielschichtigkeit und die Herausforderungen in der Evaluationspraxis zu sensibilisieren und andererseits Anregungen für einen kompetenten Umgang mit ihnen zu liefern. Zu diesem Zweck wurden die Autorinnen und Autoren dieses Bands aufgefordert, ihr Handeln gezielt auf diese kontinuierlichen Entscheidungsprozesse hin zu reflektieren. Die Beiträge sollen damit die professionellen Überlegungen von Evaluierenden offenlegen und der Leserin bzw. dem Leser die Handlungsspielräume, -alternativen und dahinterliegenden Begründungen für das jeweilige Vorgehen vermitteln.

Das Buch versteht sich damit als ein Beitrag zur Professionalisierung. Professionalisierung bezieht sich in diesem Kontext auf kompetentes, fachgerechtes Handeln von Evaluierenden und Evaluationsteams (Brand 2009, S. 26 ff.). Sie gelingt aus unserer Sicht, wenn ein fortlaufender Abgleich zwischen Praxis und Theorie stattfindet

und die praktischen Herausforderungen sowie die gefundenen Lösungen reflektiert werden.

Als Herausgeberinnen dieses Sammelbands hoffen wir, dass aus den vorliegenden Artikeln eine Bandbreite an Anregungen für zukünftige Evaluationsvorhaben gewonnen werden kann. Die Vorstellung innovativer Ansätze und Methoden soll zum eigenständigen Erproben in neuen Kontexten und Bedingungen ermutigen, zur Reflexion des vielleicht schon zur Routine gewordenen Vorgehens anregen und zur Weiterentwicklung von Ideen beitragen.

Mit diesem Band möchten wir Evaluationsinteressierte, Evaluationspraktikerinnen und -praktiker sowie Studierende, die sich mit dem Thema Evaluation befassen, ansprechen. Der Band soll sowohl Evaluationsnovizinnen und -novizen als auch erfahrenen Evaluierenden neue Erkenntnisse bieten.

Die Rolle von Univation als fachlicher Bezugspunkt

Die Herausgeberinnen dieses Buchs sind Mitarbeiterinnen von Univation, einem Institut für Evaluation, das 1997 von Wolfgang Beywl und Berthold Schobert gegründet wurde. Auch die Beiträge sind weitgehend von Personen geschrieben, die als feste oder freie Mitarbeiterinnen bzw. Mitarbeiter von Univation Evaluationen realisiert haben. Sie teilen das Evaluationsverständnis des Instituts, kennen die institutionellen Abläufe und arbeiten an verschiedenen ‚internen Projekten‘ mit, wie der Entwicklung des Glossars der wirkungsorientierten Evaluation, der Betreuung der Webseite zur Selbstevaluation, der Übersetzung von Checklisten etc. Die Herausgeberinnen und viele der Autorinnen und Autoren sind somit eng mit dem Institut Univation verbunden, weshalb die Themen und Inhalte dieses Buchs stark von der Evaluationspraxis, wie sie durch Univation gestaltet wird, geprägt sind. Aus diesem Grund soll hier kurz auf die Besonderheiten oder Spezifika des Instituts und seiner Evaluationen eingegangen werden.

Kennzeichnend für Univation ist, dass große Programmevaluationen häufig in multidisziplinären Teams durchgeführt und dabei gezielt Freiberuflerinnen und -berufler sowie Kooperationspartnerinnen und -partner einbezogen werden. Diese Teamkonstellationen sind mit einem erhöhten Abstimmungsbedarf verbunden und ziehen häufig Diskussionen um Evaluationsfachbegriffe sowie geeignete Methoden und Vorgehensweisen nach sich. Die damit einhergehende Explikation der jeweils eigenen Überlegungen und der damit verbundenen fachlichen Annahmen fördert eine professionelle Kommunikationskultur und Evaluationspraxis. Typisch für Univation ist weiterhin seine feldübergreifende Verortung. Die Evaluationsgegenstände kommen aus verschiedenen Bereichen wie der Bildungs-, Jugend- Arbeitsmarkt- und Gesundheitspolitik. Bei der Teamzusammenstellung wird darauf geachtet, dass für jede Evaluation eine fundierte Feldkompetenz gewährleistet ist.

Das Evaluationsverständnis von Univation ist stark durch seinen wissenschaftlichen Leiter, Wolfgang Beywl, geprägt. Wichtige Anregungen für die Entwicklung

seiner Evaluationsphilosophie resultieren aus Workshops, die zusammen mit Robert Stake, dem Begründer der Responsiven Evaluation (Stake 1995) und Michael Patton, Protagonist der Nutzungsfokussierten Evaluation (Patton 1997), veranstaltet wurden. In den 2000er Jahren fand – nun gemeinsam mit dem Kernteam von Univation – eine intensive Auseinandersetzung mit Verfahren zur Wirkungsmodellierung statt. Daraus ging der Programmbaum als Instrument für die Evaluationsplanung und -umsetzung hervor (Beywl/Niestroj 2009). Zudem wurden Stärken und Schwächen theoriebasierter Ansätze untersucht (Giel 2013). Die kontinuierliche Beschäftigung mit verschiedenen Evaluationsmodellen zieht sich durch die Institutsgeschichte (Beywl 1988, Beywl/Speer/Kehr 2003, Beywl 2006).

Ein weiteres Anliegen von Univation ist es, die fachliche Reflexion über Evaluation zu fördern und damit zu einer Weiterentwicklung der Evaluationskultur beizutragen. Dies geschieht über die Reflexion der eigenen Praxis, die Arbeit an einer geteilten Evaluationsfachsprache (insbesondere in Form des Glossars der wirkungsorientierten Evaluation), den fachlichen Austausch innerhalb der Gesellschaft für Evaluation, die Verbreitung der Standards für Evaluation (Widmer/Beywl 2000) und die Umsetzung von Evaluationsforschung (Beywl/Speer 2004, Giel 2013, Mäder 2013).

Zudem leistete Univation von Beginn an einen Beitrag dazu, Evaluationswissen und -kompetenzen bspw. mit Fachpublikationen und Handreichungen praxisorientiert zu vermitteln (Beywl u. a. 2007, Beywl/Schepp-Winter 2000, Keller-Ebert/Kißler/Schobert 2005, Farrokzhad/Mäder 2014) und so insbesondere die Weiterbildung von Praktikerinnen und Praktikern sowie Personen, die Evaluationen in Auftrag geben, zu unterstützen. Längere Zeit lag dabei ein Schwerpunkt auch auf der Weiterbildung im Bereich der Selbstevaluation (Beywl/Bestvater/Friedrich 2011). Als langjähriger Leiter der Weiterbildungsstudiengänge zur Evaluation des Zentrums für universitäre Weiterbildung der Universität Bern baute Wolfgang Beywl ein umfangreiches Qualifizierungsangebot auf. Ergänzend wird nun auch der Aufbau von internem Evaluationsvermögen in den Blick genommen, verbunden mit einem darauf abgestimmten Weiterbildungskonzept (Balzer/Beywl 2015).

Zum Aufbau des Buches

Der vorliegende Band präsentiert Aufsätze aus drei inhaltlichen Bereichen: Professionalisierungsinstrumente, Evaluationsansätze und Evaluationsmethoden. Nachfolgend werden die drei inhaltlichen Schwerpunkte kurz vorgestellt.

Professionalisierungsinstrumente: In den Handbüchern und Lehrbüchern zur Evaluation liegt der Schwerpunkt meist auf theoretischen, methodologischen und methodischen Fragen. Über die professionelle Planung, Organisation und Umsetzung von Evaluationen wird weniger berichtet. In der Praxis zeigt sich jedoch, dass Aspekte der Steuerung und die damit verbundene Professionalisierung wie die Auswahl des Personals, die Arbeitsformen und -organisation sowie die Kommunikation

untereinander für den Projekterfolg ebenso entscheidend sind. Oft werden Evaluationen von einer Gruppe von Evaluierenden geplant und durchgeführt, was für die Steuerung komplexer Programmevaluationen immer wieder neue Herausforderungen mit sich bringt. Wir haben dieses Thema als eigenen inhaltlichen Bereich an den Anfang dieses Bands gestellt, um seine Bedeutung für die Durchführung professioneller Evaluation zu betonen. Dabei verstehen wir unter Steuerung in einem weiten Sinn nicht nur das konkrete Management einzelner Evaluationen, sondern weitergehend auch Quellen, Ansätze, Instrumente für die Professionalisierung der Steuerung von Evaluationen. Hierzu zählen aus unserer Sicht u. a. die Potenziale einer konsistenten Begriffsverwendung, die Orientierung an geprüften und hochwertigen Arbeitshilfen wie Checklisten, Strategien zur Förderung des Datenschutzes, die kreative Nutzung heterogener Kompetenzen in multiprofessionellen Teams und die Chancen einer interkulturellen Öffnung von Evaluationsteams.

Evaluationsansätze: Die in diesem Band vorgestellten Evaluationsansätze haben in besonderem Maße den Nutzen und die Nutzung von Evaluationen im Blick. Sie intendieren in erster Linie, die betrachtete Praxis weiterzuentwickeln. Dabei denken sie Evaluation von einer komplexen und sich entwickelnden Programmrealität aus. Gerade der von Univation entwickelte Programmbaum bietet einen Zugang zur Komplexität und Veränderlichkeit von Evaluationsgegenständen. Zudem eignen sich die vorgestellten Evaluationsansätze für die Evaluation von experimentierenden und unausgereiften Programmen. Unserer Ansicht nach stellt diese Art von Programmen die Evaluation vor besondere Herausforderungen, wie sich schnell verändernde Kontexte und Programmumsetzungen, nicht explizierte Wirkannahmen, eine Vielzahl an Programmebenen und Projekten. Nützliche Lösungen bieten hier aus unserer Sicht insbesondere die Developmental Evaluation sowie die Wirkungsmodellierung in Rahmen theoriebasierter Evaluationsansätze.

Evaluationsmethoden: Der dritte Teil des Bands stellt Evaluationsmethoden in der Praxis vor. Bei diesem Thema war es uns wichtig, neben klassischen Datenerhebungs- und Auswertungsmethoden besonders auch für die Phase der Gegenstandsbestimmung sowie die der Ableitung von Schlussfolgerungen und Bewertungen methodische Herangehensweisen vorzustellen. Aus unserer Sicht sollte der gesamte Evaluationsprozess methodisch gerahmt sein. Aufholbedarf zeigt sich dabei vor allem für die Phasen der Gegenstandsbestimmung und Ergebnisvermittlung. Wir möchten daher den Teil zu Evaluationsmethoden dafür nutzen, zusätzlich zu neuen Erkenntnissen im Rahmen der Datenerhebung und -auswertung insbesondere auch Verfahren für die methodische Bearbeitung dieser Phasen vorzustellen. So wird für die Phase der Gegenstandsbestimmung präsentiert, wie Ziele so entwickelt werden können, dass sie für die Programmsteuerung und Evaluation nützlich sind. Weiterhin werden für die Bewertung des Programmerfolgs partizipative methodische Verfahrensweisen präsentiert. Für den Bereich der Datenerhebung und -auswertung wurden Beiträge gewählt, die Antworten auf spezielle Herausforderungen liefern, wie hochkomplexe Programme oder schwierig zugängliche Zielgruppen. Nicht zuletzt beinhaltet der Methodenabschnitt einen Beitrag zur Analyse und Interpre-

tation qualitativer Daten wie auch der Einbezug webbasierter Technologien für die Informationsbeschaffung, Datenauswertung und Ergebnispräsentation.

Im Folgenden werden die einzelnen Artikel in den drei inhaltlichen Bereichen kurz präsentiert. Im ersten Teil „Professionalisierungsinstrumente“ finden sich fünf Beiträge.

Den Anfang machen *Melanie Niestroj*, *Wolfgang Beywl* und *Berthold Schobert*, die Chancen der Verwendung eindeutig definierter Begriffe anhand eines Glossars (Eval-Wiki: Glossar der Evaluation) präsentieren. Im Artikel wird dargestellt, wie das Eval-Wiki entwickelt wurde, wieso ein konsistentes Vokabular in der Evaluation wichtig ist und welchen Beitrag das Glossar zur weiteren Professionalisierung leisten kann.

Im Anschluss präsentiert *Katharina Klockgether*, ausgehend von dem Evaluation Checklist Project der Western Michigan University, verschiedene Checklisten sowie deren Anwendungsspektrum. Sie reflektiert deren Nützlichkeit für den Evaluationsalltag sowie die Risiken ihrer Nutzung. Anhand einer Literaturrecherche und einer Umfrage in der Evaluationscommunity wird über die Verbreitung und Nutzung von Checklisten in der deutschsprachigen Evaluation berichtet.

Marc Jelitto widmet sich mit dem Datenschutz einem für Evaluationen besonders brisanten Thema. Er klärt darüber auf, welche Gefahren die Datensicherheit insbesondere in Evaluationen bedrohen, und stellt Schutzmaßnahmen vor. Dabei werden viele Beispiele aus der Praxis aufgezeigt und praktische Anwendungshinweise gegeben, bspw. zur Speicherung und Versendung personenbezogener Daten.

Der Artikel von *Schahrzad Farrokhzad* beschäftigt sich mit interkulturellen Kompetenzen in Evaluationen. Sie unterstreicht die Wichtigkeit einer interkulturellen Öffnung und stellt verschiedene Strategien vor, wie diese im Rahmen von Evaluationsprojekten angegangen werden kann. Die Autorin stellt ein Kompetenzmodell vor, das für die Ausbildung interkultureller Kompetenzen bei Evaluierenden genutzt werden kann.

Im letzten Artikel in diesem ersten Bereich beschreiben *Dörte Schott* und *Wolfgang Beywl* den Steuerungsansatz der Schwarmevaluation. Er zeigt, wie durch die Kollaboration mehrerer Evaluierender im Rahmen von parallelen Mikro-Evaluationen Effizienzgewinne und eine hohe Nützlichkeit erzielt werden. Die Prinzipien einer Schwarmevaluation werden anhand eines Fallbeispiels aufgezeigt und förderliche Bedingungen für ihre Umsetzung präsentiert.

Im zweiten Abschnitt des Bandes geht es um Evaluationsansätze.

Samera Bartsch, *Wolfgang Beywl* und *Melanie Niestroj* stellen den „Programmbaum“ von Univation vor; ein logisches Modell, das Programme wirkungsorientiert abbildet und Evaluationen konzeptionell strukturieren kann. Der Programmbaum stellt keinen eigenen Evaluationansatz dar, sondern einen unterstützenden Schlüsselbestandteil verschiedener Evaluationsansätze. Die Entstehungsgeschichte des Programmbaums wird aufgezeigt und seine Elemente werden erläutert. Der Beitrag

möchte deutlich machen, wie der Programmbaum dabei hilft, den Evaluationsgegenstand systematisch darzustellen und so besser verständlich und untersuchbar zu machen.

Susanne Giel geht der Frage nach, wie Evaluationen Wirkungen von Programmen erfassen können, wenn klassische Wirkungsforschungsdesigns nicht anwendbar sind. Hierzu schlägt sie das Konzept einer theoriebasierten Evaluation vor, die Programmtheorien in den Mittelpunkt stellt. Die praktische Umsetzung wird anhand der Evaluation von Modellprogrammen zur Prävention von Rechtsextremismus, Antisemitismus und Rassismus illustriert.

Maria Gutknecht-Gmeiner erläutert das von Michael Q. Patton publizierte Konzept der Developmental Evaluation, das eine Rahmung für Evaluationen bietet, welche die Aufgabe haben, Programme in ständiger Entwicklung aktiv zu begleiten. Sie reflektiert die Potenziale dieses Vorgehens am Beispiel der Evaluation eines innovativen Ausbildungsprogramms, das den Zugang für Mädchen und junge Frauen zu technischen Berufen ebnet soll.

Im dritten Teil des Buches finden sich sieben Beiträge zu Evaluationsmethoden. *Stefan Schmidt* präsentiert eine wirkungsorientierte Zielsystematik, die sowohl für die Steuerung der Evaluation als auch für die Programmsteuerung Nutzen erzeugt. Er entwickelt Anforderungen an eine Ordnung von Zielen und deren Formulierung und beschreibt diese anhand ihres praktischen Einsatzes in einem Bildungsnetzwerk.

Katharina Klockgether beschäftigt sich in ihrem Artikel mit den Chancen und Risiken von Monitorings in der Evaluation. Anhand von zwei verschiedenen Monitoring-Instrumenten aus einem Demokratie- bzw. Arbeitsmarktprogramm werden relevante Schritte bei der Planung, Implementierung und Auswertung der jeweiligen Instrumente veranschaulicht. Zudem werden bestehende Monitorings aus den Bereichen Gesundheit, Arbeitsmarkt, Bildung und Landwirtschaft vorgestellt, auf die Evaluierende zurückgreifen können.

Einen Einblick in die Durchführung von Face-to-Face-Interviews mit Jugendlichen liefert der Beitrag von *Matthias Sperling*. Er zeigt auf, welche Potenziale und Herausforderungen mit der Befragung von Jugendlichen in Evaluationen verbunden sind und wie diesen begegnet werden kann. Thematisiert werden entwicklungspsychologische Besonderheiten von Jugendlichen, ihre Sprache und ihr Verhalten in Interviewsituationen. Davon ausgehend werden Hinweise für zentrale Aufgaben wie die Gestaltung der Interviewatmosphäre oder die Leitfadententwicklung gegeben.

Die Potenziale von Gruppendiskussionen für die Ermittlung von implizitem Wissen beschreibt der Aufsatz von *Ute B. Schröder*. Im Fokus stehen die Rekonstruktion von kollektiven Orientierungen mittels der dokumentarischen Methode und deren Bedeutung für die Evaluation. Gezeigt wird, wie Änderungen von Einstellungen und Werthaltungen bei den Zielgruppen von Programmen durch eine sorgfältige Analyse des Diskurses erfasst werden können.

Marianne Lück-Filsinger arbeitet den Nutzen der Grounded Theory für die Konzeption und Auswertung von Erhebungen in der Evaluation heraus. Ihr Beitrag fokussiert auf die Produktivität dieses Forschungsstils und seiner Kodierverfahren für qualitative Fragestellungen in Evaluationen. Anhand eines Pilotprojekts zur Übergangsgestaltung zwischen Kindergarten und Grundschule werden die Schritte des Offenen, Axialen und Selektiven Kodierens veranschaulicht.

Im Beitrag ‚Evaluation und das Web‘ präsentiert *Elitsa Uzunova* die Chancen von Internettechnologien für zentrale Evaluationsaufgaben. Sie beschreibt neue Möglichkeiten für die Datengenerierung, -erhebung und -auswertung sowie die Präsentation von Ergebnissen. Zudem zeigt sie den Nutzen webbasierter Lösungen für den Wissenstransfer in der Evaluationscommunity auf.

Susanne Mäder geht in ihrem Aufsatz der Frage nach, wie Evaluationen zu Bewertungen gelangen können. Nach einem kurzen Überblick über den Umgang mit Bewertung in Evaluationstheorie und -praxis präsentiert sie zwei Bewertungsverfahren, die in der Evaluation von Modellausbildungen in der beruflichen Bildung entwickelt wurden. Im ersten Beispiel dienen Wirkmodelle als Ausgangspunkte für Bewertungskriterien. Im zweiten Beispiel steht die partizipative Festlegung von Erfolgspunkten im Vordergrund.

Hinweis für Leserinnen und Leser

Klarheit und Transparenz in der Evaluationsfachsprache sind aus Sicht der Herausgeberinnen zentrale Voraussetzungen für die produktive Verständigung über Herausforderungen und Lösungen in der Evaluationspraxis. Uns ist es daher ein Anliegen, eine eindeutige und geteilte Definition von Evaluationsbegriffen zu fördern. Aus diesem Grund wurde bei der Redaktion und dem Lektorat der eingereichten Beiträge auf eine konsistente Nutzung zentraler Evaluationstermini geachtet. Als Bezugspunkt für die einheitliche Begriffsverwendung wurde das Glossar der wirkungsorientierten Evaluation von Univation genutzt, das über 400 Begriffe und ihre Definitionen enthält. Das Glossar wurde erstmals Ende 2004 als Printmedium veröffentlicht und ist seit 2010 auch online kostenfrei unter www.eval-wiki.org/glossar zugänglich. Die Begriffsdefinitionen des Glossars bilden eine verbindliche Grundlage für die Beiträge in diesem Band. Für jeden einzelnen Beitrag sind jeweils die im Eval-Wiki definierten Begriffe bei ihrer jeweils ersten Nennung markiert: In der Printausgabe sind sie grau gedruckt, in der E-Book-Ausgabe sind die Begriffe mit dem Eval-Wiki direkt verlinkt.

Wir möchten uns an dieser Stelle bei Kolleginnen und Kollegen für ihre tatkräftige Unterstützung bedanken: bei Jeyhun Alizade für die technische Umsetzung der Links zu den Glossarbegriffen, bei Melanie Niestroj sowohl für die Identifizierung der Glossarbegriffe in den Artikeln und die Überprüfung ihrer definitionsgemäßen Verwendung als auch für viele wichtige inhaltliche Hinweise und bei Wolfgang Beywl für die vielen inspirierenden Anregungen.

Beitrag zu einer konsistenten Evaluationsterminologie – das Glossar von Univation

Melanie Niestroj/Wolfgang Beywl/Berthold Schobert

1. Babylon ist überall

Dokumente (über-)staatlicher Organisationen oder Qualitätshandbücher von Unternehmen enthalten zahlreiche für Evaluationen wichtige Begriffe. Wenn es z. B. darum geht zu bezeichnen, was ein operationalisiert festgelegter Zustand ist, der durch ein Programm ausgelöst werden soll, so wird dieser Zustand wie folgt oder anders bezeichnet: Soll-, Kenn-, Mess- oder Zielgröße, Qualitäts- oder Erfolgskriterium, Ziel- oder Schwellenwert, (Mindest-)Standard, Rubric, Indikator. Dabei werden die Begriffe – oft sogar im selben Dokument – synonym und nicht selten auch noch zur Bezeichnung ganz anderer Sachverhalte verwendet. Schließlich werden sogar Sätze formuliert, welche diese an anderen Textstellen synonym verwendeten Bezeichnungen miteinander in unterschiedliche Verbindungen bringen, als Ober- und Unterbegriff, Veranschaulichung o. Ä. Schließlich werden die Bezeichnungen durch andere spezifiziert, z. B. als ‚qualitative‘ und ‚quantitative‘ Indikatoren, Ziele, Messgrößen usw., und es bleibt der Intuition der Lesenden überlassen, was darunter zu verstehen ist, da auch diese Attribute nirgends definiert sind. Es fällt überdies auf, dass gelegentlich selbst für zentrale Begriffe – die z. B. in der Überschrift eines Grundlagenpapiers oder dem Titel eines Buches genannt sind – keine explizite Definition gegeben wird.

Während eine solche babylonische Sprachverwirrung auf der Ebene von Programmen alleine bereits verstören kann, gewinnt sie an Brisanz, sobald eine Evaluation hinzukommt, die bspw. über das Programm berichtet. Dabei entsteht die zusätzliche Anforderung auseinanderzuhalten, ob mit einem Begriff ein Sachverhalt auf der Ebene des Programms (dem Untersuchten) oder der Evaluation (der Untersuchung) angesprochen ist. Ein typisches Beispiel: Werden im Englischen zwei ganz unterschiedliche und klar zugeordnete Begriffe, nämlich ‚results‘ (of the program) und ‚findings‘ (of the evaluation), genutzt, sprechen deutschsprachige Evaluationsberichte nicht selten lediglich von ‚Ergebnissen‘. Um die Interpretation zu erleichtern haben wir die Konvention eingeführt, zum einen von ‚Resultaten‘ (eines Programms) und zum anderen von ‚Ergebnissen‘ (einer Evaluation) zu sprechen und zu schreiben. Wir versuchen diese wie zahlreiche andere fachsprachliche Konventionen, in unseren Texten konsequent umzusetzen. Unterstützt werden wir dabei durch das ‚Eval-Wiki: Glossar der Evaluation‘ (www.eval-wiki.org/glossar), ein Online-Glossar, in dem aktuell 424 evaluationsrelevante Begriffe definiert sind und das allen Interessierten kostenfrei zur Verfügung steht.

Checklisten zur Planung und Steuerung von Evaluationen

Katharina Klockgether

Häufig wird von unerfahrenen Evaluierenden der Wunsch nach einem „Rezept“ für die Planung und Durchführung einer Evaluation geäußert. So etwas gibt es nicht. Was es jedoch gibt – bisher in der deutschsprachigen Evaluations-Community noch nahezu unbeachtet – sind evaluationsspezifische Checklisten.

Checklisten retten bei Einsatz in der Passagierluftfahrt regelmäßig Menschenleben. Sie helfen Piloten und Pilotinnen in seltenen, gefährlichen Notsituationen die richtigen Entscheidungen zu treffen und lebensrettende Schritte nicht zu vergessen (Gawande 2009). Wie Checklisten in der Evaluation hingegen zwar nicht Leben retten, aber durchaus die Möglichkeit bieten, die Qualität von Evaluationsvorhaben zu steigern, welche Schwierigkeiten dabei auftreten können und welche Tücken beachtet werden müssen, soll dieser Artikel verdeutlichen. Durch eine Kooperation der Western Michigan University (WMU) und Univation sind einige der unten beschriebenen Evaluations-Checklisten, welche auf der Webseite der WMU veröffentlicht sind,¹ auch bei Univation im regelmäßigen Gebrauch. Die Erfahrung zeigt, dass die Verwendung von Evaluations-Checklisten eine systematische Herangehensweise bietet, um einerseits Schritte der Evaluation zu planen und zu prüfen und um andererseits als Bewertungsinstrument die Güte, Bedeutsamkeit und Tauglichkeit eines Evaluationsgegenstands, bspw. eines Programms, zu bestimmen.

1. Welche Arten von Checklisten gibt es?

In der Literatur werden Checklisten allgemein als Erinnerungshilfe umschrieben. Sie bezeichnen eine systematische Sammlung von Faktoren, welche berücksichtigt werden müssen, um eine Aufgabe gut erledigen zu können bzw. ein bestimmtes Ziel zu erreichen (Scriven 2000). Checklisten liefern Hinweise, was bei der Planung, Ausführung und Berichterstattung beachtet werden muss. Sie listen auf, was beachtet werden muss, wenn ein bestimmtes Evaluationsmodell angewendet werden soll (bspw. Utilization Focused Evaluation) oder unterstützen bei der Bewertung des Evaluationsgegenstands. In der Evaluation können Checklisten von einzelnen Evaluierenden zur Vorbereitung und Durchführung von Evaluationen genutzt werden (formative Orientierung). Es kann jedoch auch im Team damit gearbeitet werden. Eine weitere Gruppe von potenziellen Nutzenden sind Personen oder Institutionen, die eine Evaluation in Auftrag gegeben haben und die Güte der Evaluation bewer-

1 Die Checklisten sind hier abrufbar: <http://www.wmich.edu/evaluation/checklists> [Stand: 19.06.2015]. Deutsche Übersetzungen achtzehn dieser Listen können hier abgerufen werden: <http://www.univation.org/checklisten> [Stand: 19.06.2015]

Informationssicherheit in der Evaluation

Marc Jelitto

Laut den Standards der DeGEval Gesellschaft für Evaluation e. V. sollen „Evaluationen so geplant und durchgeführt werden, dass Sicherheit, Würde und Rechte der in eine Evaluation einbezogenen Personen geschützt werden“ (Abschnitt Fairness F2: Schutz individueller Rechte). Dazu gehört auch der Schutz erhobener Daten der befragten Personen und aller Beteiligten einer Evaluation. Alle *Evaluationsmitarbeitenden* sollten daher das Themenfeld *Datenschutz* kennen und in ihrer Arbeit berücksichtigen. In Deutschland hat der Datenschutz eine langjährige Geschichte, die Gesetze schreiben einen sorgfältigen Umgang mit personenbezogenen Daten vor. Namen von Stakeholdern, deren Adressdaten und in der Evaluation gewonnene Daten werden sicher verwahrt und landen z. B. nicht im Altpapier, sondern im Reißwolf bzw. werden separat sicher entsorgt. Bei Umfragen werden Ergebnisse von Nutzerdaten getrennt erhoben bzw. gespeichert, Details werden anonymisiert oder so umformuliert, dass eine interviewte Person im Endbericht nicht wieder erkannt werden kann. Bei Online-Erhebungen werden IP-Adressen von Webseitenbesuchern nur verkürzt abgespeichert und ggf. auf den Einsatz von Webanalyse-Werkzeugen hingewiesen. Zu einem solchen Vorgehen sind Evaluierende gesetzlich und ethisch verpflichtet. Allerdings arbeiten sie neben Daten in einem engen Sinne mit weiteren Informationen, die einen hohen Schutzbedarf haben. Daher widmet sich dieser Beitrag dem *breiter gefassten Begriff Informationssicherheit*, also der sensiblen Behandlung aller schützenswerten Informationen.

Während im zwanzigsten Jahrhundert Gefahren für Informationsdiebstähle in der Regel eine körperliche Anwesenheit vor Ort (wie durch einen Einbruch) erforderten, schafft die immer intensivere Computer-Vernetzung sowie die verstärkte Nutzung von Online-Diensten zum Datenaustausch neue Gefahren. Hinzu kommt, dass Irrtum und Nachlässigkeit am PC größere Auswirkungen haben können, wie z. B. das Löschen eines falschen Ordners. Nach dem ‚Bericht über die Evaluation des Bundesgesetzes über den Datenschutz‘ des Schweizerischen Bundesrates (2011, S. 350) sollten „betroffene Personen [] verstärkt für die mit technologischen Entwicklungen einhergehenden Risiken sensibilisiert werden“. Zu den Betroffenen gehören auch Evaluatorinnen und Evaluatoren.

Dieser Buchbeitrag soll im Evaluationsbereich für eine ähnliche Sensibilisierung beim Thema Informationssicherheit sorgen, wie sie beispielsweise durch eine Initiative der Bundesakademie für öffentliche Verwaltung (BAKöV) vorbildlich initiiert wurde: 2009 bis 2012 führte die BAKöV Schulungen für ca. 80.000 Bundesbedienstete durch. Die dreistündigen Sensibilisierungsmaßnahmen mit Vorträgen, Videos und Kurzdiskussionen wurden bei über 150 Behörden wie Ministerien, Hauptzollämtern usw. durchgeführt. Live-Hacking-Veranstaltungen, in denen vor den Augen

Interkulturelle Kompetenz – ein weiterer Baustein für das professionelle Handeln in Evaluationen

Schahrzad Farrokhzad

Interkulturelle Kompetenzen werden seit geraumer Zeit von verschiedensten Berufsgruppen gefordert. Seit Jahren werden interkulturelle und diversitätsorientierte Fortbildungen nachgefragt und angeboten, u. a. für den Bildungs- und Sozialbereich, die öffentliche Verwaltung, die Entwicklungszusammenarbeit, aber auch für Wirtschaftsunternehmen. Gleichzeitig haben in der Bundesrepublik, oft gefördert von Bund, Ländern, Kommunen oder Stiftungen, Modellprogramme und -projekte rund um die Themen Migration, Integration, Vielfalt und Interkulturelle Öffnung Hochkonjunktur. Entsprechend nehmen die zu diesen Themenfeldern ausgeschriebenen Evaluationen weiter zu.

Doch wie steht es eigentlich um die Interkulturellen Kompetenzen von Evaluatoredinnen und Evaluatoren? Sind solche Kompetenzen nicht nur für die Akteursgruppen der praktischen Projektarbeit in diesen Feldern, sondern auch für professionell Evaluierende nützlich und sinnvoll? In diesem Beitrag wird der Frage nachgegangen, ob und inwiefern Interkulturelle Kompetenzen in Evaluationen von Vorteil sein können. Entlang eines Modells Interkultureller Kompetenzentwicklung werden verschiedene Dimensionen Interkultureller Kompetenz und ihre Bedeutung für die professionelle Evaluationspraxis behandelt. Dies wird anhand von Beispielen aus verschiedenen Evaluationsprojekten veranschaulicht.

Kompetenzanforderungen an Evaluatoredinnen und Evaluatoren – Vorbemerkungen

Welche Kompetenzen sind für professionelle Evaluationen notwendig und wünschenswert? Die Professionalisierung von Evaluationen ist weiterhin ein aktuelles und zentrales Thema. Die Gesellschaft für Evaluation (DeGEval) widmete im Jahre 2014 ihre Jahrestagung unter dem Titel „Professionalisierung in und für Evaluationen“ diesem Thema. Hierbei ging es u. a. um die Standards für Evaluationen¹, die Qualität von Aus- und Fortbildungen im Bereich der Evaluationen und deren Zertifizierung. Darüber hinaus wurde das Selbstverständnis von Evaluatoredinnen und Evaluatoren als Akteure im Kontext gesellschaftlicher Entwicklungen verhandelt.

In den von der Gesellschaft für Evaluation herausgegebenen „Empfehlungen für die Aus- und Weiterbildung in der Evaluation“ (DeGEval 2008a, S. 9 ff.) wird

1 http://www.degeval.de/images/stories/Publikationen/DeGEval_-_Standards.pdf [Stand: 18.03.2015]

Schwarm-Evaluation als Steuerungsansatz

Dörte Schott und Wolfgang Beywl

„... our world is becoming so complex that it cannot be comprehended by any single human being. Swarm intelligence offers an alternative way of designing ,intelligent‘ systems in which autonomy, emergence, and the ability to distribute tasks replace control, preprogramming, and centralization ...“ (Gloor 2006, S. 21)

In den letzten Jahren sind steigende Ansprüche an geförderte – und zu evaluierende – Programme mit zunehmend komplexen Programmstrukturen bei zugleich zunehmendem Kostendruck zu beobachten. Sowohl der Kostendruck als auch die gestiegenen Ansprüche werden teilweise an die Anbieter von Evaluationen weitergegeben. Große, zentral aufgesetzte Förderprogramme für eine Vielzahl paralleler Projekte, häufig mit innovativer Zielrichtung, haben nicht selten ein immenses finanzielles Gesamtvolumen, von dem aber oft nur ein sehr geringer Anteil auf das einzelne Projekt – und damit auch auf die Evaluation dieses Projekts – entfällt. Für die Evaluation solcher Projekte wird in diesem Beitrag ein Ansatz zur Steuerung vorgestellt: der Schwarm-Evaluations-Ansatz (SEA).

Der Steuerungsansatz ist vor dem Hintergrund von 14 parallelen small-scale Evaluationen (vgl. Abschnitt 2.2) entstanden, welche zwischen 2005 und 2008 für Entwicklungspartnerschaften der zweiten Europäischen EQUAL-Initiative (vgl. Abschnitt 2.1) durch oder in Kooperation mit Univation Institut für Evaluation Dr. Beywl & Associates durchgeführt wurden.

Als besonders herausfordernd stellten sich dabei die hohen Anforderungen an die Evaluation bei gleichzeitig knapper Budgetierung dar. Die Anforderungen ergaben sich zum einen aus den Leitfäden des Programms EQUAL und zum anderen aus der Komplexität und Emergenz der Evaluationsgegenstände. Beispielsweise sollten im vorliegenden Fall Netzwerke aufgebaut werden, was bei Erfolg zahlreiche Interaktionsbeziehungen zwischen den beteiligten Teilprojekten nach sich zieht. Dies erhöht für die Evaluation nicht nur den Aufwand im Rahmen der Datenerhebungen, sondern auch jenen für evaluationsbegleitende Abstimmungs- und Aushandlungsprozesse. Eine weitere Herausforderung bestand darin, den Projekterfolg zu definieren, da aus den Netzwerken innovative Lösungen für den Arbeitsmarkt hervorgehen sollten, welche bei Projektstart noch weitgehend unklar oder unbekannt waren. Für alle Evaluierenden stellte sich entsprechend die Frage: Wie kann unter diesen Bedingungen eine fachlich gute Evaluation mit einer hohen Nützlichkeit für die Stakeholder aussehen?

Gesucht war daher ein Kreativität und Effizienz fördernder Steuerungsansatz für Evaluation. Derzeit bieten das Internet und die moderne Kommunikationstechnologie günstige Voraussetzungen für kollaborative Kreativität und effiziente Zusammenarbeit durch globale und zeitnahe Kommunikationsmöglichkeiten (vgl.

Der Programmbaum als Evaluationsinstrument¹

Samera Bartsch/Wolfgang Beywl/Melanie Niestroj

Der von Univation entwickelte Programmbaum hat sich ob seiner Anschaulichkeit und ‚didaktischen Anleitungskraft‘ zu einem weit verbreiteten Hilfsmittel für die Planung und Durchführung von Evaluationen entwickelt. Er wird in zahlreichen Feldern nicht nur von Evaluierenden, sondern auch von Programmverantwortlichen und -mitarbeitenden angewendet.² Dieser Beitrag stellt den Programmbaum vor und erläutert ihn anhand von Beispielen. Seine möglichen Beiträge für die Arbeit der Programmverantwortlichen wie der Evaluationsverantwortlichen werden dargestellt. Abschließend wird auf Grenzen der Nutzung des Programmbaums eingegangen.

1. Zielsetzung des Beitrags

Zu Beginn einer Evaluation, oft sogar schon früher, bei der Angebotserstellung, bietet sich eine teils überraschend schwierige Aufgabe: den Gegenstand der geplanten Evaluation, also das Programm, welches evaluiert werden soll, zu verstehen und zu beschreiben. Fragen wie folgende sind zu klären: Aus welchen Elementen besteht das Programm und wie sind diese miteinander verbunden? Von wem wird das Pro-

-
- 1 Dieser Beitrag basiert auf folgenden bereits publizierten Beiträgen: Beywl, Wolfgang (2006): Demokratie braucht wirkungsorientierte Evaluation – Entwicklungspfade im Kontext der Kinder- und Jugendhilfe. In: Projekt eXe (Hrsg.): Wirkungsevaluation in der Kinder- und Jugendhilfe. Einblicke in die Evaluationspraxis, S. 25–46; und Beywl, Wolfgang/Niestroj, Melanie (2009): Der Programmbaum. Landmarke wirkungsorientierter Evaluation. In: Univation (Hrsg.): Das A-B-C der wirkungsorientierten Evaluation. 2. Aufl., Köln: Univation – Institution für Evaluation GmbH, S. 137–149.
 - 2 Eine Google-Abfrage mit den Stichworten ‚Programmbaum‘ und ‚Evaluation‘ zeigt neben zahlreichen Publikationen und Tagungsdokumentationen von Univation und Kooperationspartnern, dass auch weitere Einrichtungen die Verwendung des Programmbaums zur Planung einer Evaluation empfehlen. Neben den Webseiten von öffentlichen Verwaltungen finden sich auch auf den Webseiten von Organisationsberatungen Erläuterungen zum Programmbaum. Siehe bspw. <http://www.impulse.at/pages/deutsch/evaluation/programmtheorien-logische-modelle.php> [Stand: 03.03.2015] oder http://www.systemblick.de/uploads/media/Evaluation-und-Erfolgskontrolle_systemblick.pdf [Stand: 03.03.2015]. Ein prominentes Beispiel bietet das Schweizerische Bundesamt für Gesundheit. Im Rahmen der Arbeit des Tabakpräventionsfonds wird der Programmbaum als Instrument der wirkungsorientierten Projektplanung empfohlen (siehe http://www.bag.admin.ch/tabak_praevention/07165/07167/index.html?lang=de) [Stand: 03.03.2015]. Neuerdings wurde der Programmbaum auch für eine Meta-Evaluation eingesetzt (Michel 2015; s. u. im Text).

Wirkungen auf der Spur mit Programmtheorien

Susanne Giel

Oftmals steht die Feststellung von Wirkungen im Zentrum von Evaluationsaufträgen. Traditionelle Wirkungsforschungsdesigns sind allerdings im Rahmen von Evaluationen häufig nicht realisierbar. Im Folgenden wird ein theoriebasierter Evaluationsansatz vorgestellt, der einen alternativen Zugang zu Wirkzusammenhängen ermöglicht. Theoriebasierte Evaluationen zeichnen sich dadurch aus, dass sogenannte Programmtheorien die zentrale Evaluationsgrundlage bilden. Der Beitrag wird zunächst umreißen, worin die Schwierigkeiten klassischer Wirkungsforschung liegen. Anschließend erfolgt eine Erläuterung, was sich hinter dem Schlüsselbegriff Programmtheorie verbirgt und wie diese für die Feststellung von Wirkungen genutzt werden kann. Diese Nutzung von Programmtheorien wird entlang der Evaluation von Modellprogrammen in den vom Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) aufgelegten Programmen *Vielfalt tut gut*, *Jugend für Vielfalt, Toleranz und Demokratie* und *Toleranz fördern – Kompetenz stärken* beispielhaft illustriert.

1 Die Herausforderung der Wirkungsfeststellung

Aussagen zu Wirkungen und der Wirkfähigkeit¹ von Programmen, Projekten und Maßnahmen (im Folgenden wird zur Vereinfachung nur von Programmen gesprochen) zu treffen, ist eine Herausforderung, die Evaluationen beschäftigt, solange es Evaluationen gibt. Dabei interessieren sich Politik – ‚Bewirken die von uns aufgelegten und finanzierten Programme das, was sie bewirken sollen?‘ – gleichermaßen wie Praxis – ‚Was bewirkt unsere Arbeit?‘ – und Gesellschaft – ‚Welche Wirkung haben Reformen oder einzelne Angebote?‘ – für Aussagen zur Wirkung von Programmen.

Für die Durchführenden von Evaluationen stellt sich die Frage, wie sie sich einen Zugang zu Wirkungen verschaffen können. Die ideale Lösung wäre eine Zeitmaschine, mit Hilfe derer sich feststellen ließe, was mit und was ohne Programm geschehen wäre. Da dieses Instrument in das Reich der Utopie gehört, ist *jeder* Zugang zu Wirkungen als eine Annäherung an eine eindeutige Wirkungsfeststellung zu begreifen. Traditionell orientieren sich Evaluationen an experimentellen Designs, wie eben auch die klassische Wirkungsforschung. Deren Grundprinzip besteht darin, ein Vergleichsszenarium zu schaffen, in dem versucht wird, alle möglichen Einflussfaktoren zwischen Ursache (Programm) und der daraus resultierenden Wirkung zu kontrollieren, um damit eindeutig den Einfluss der Intervention auf das Resultat

1 Im Unterschied zu einem empirischen Wirkungsnachweis bezeichnet der Begriff Wirkfähigkeit das Potenzial von Interventionen, spezifische Resultate zu erzielen.

Developmental Evaluation nach Michael Patton

Begriffsbestimmung und Reflexion der praktischen Anwendung

Maria Gutknecht-Gmeiner

Inspiration usually comes during work, rather than before it.
Madeleine L'Engle

Innovative Projekte entwickeln Neues. Was dieses ‚Neue‘ genau ist, ergibt sich oft erst im Tun bzw. wird Schritt für Schritt ergründet; allgemein gehaltene Zielvorstellungen werden im Verlauf der Durchführung mit Inhalten gefüllt. Der Ausgang kann für gewöhnlich nicht mit Sicherheit vorhergesagt werden, zumal Innovationen meist in Kontexten angesiedelt sind, die selbst wenig stabil sind und in denen viele Faktoren wirksam werden. Wichtige Betroffene und Beteiligte, v. a. offizielle Stellen und öffentliche Geldgeber, die ihrerseits Rechenschaft ablegen müssen, stehen solchen Projekten aufgrund der Unwägbarkeiten oft skeptisch gegenüber.

Innovative Maßnahmen stellen aus diesen Gründen nicht nur für die Projektteams, sondern auch für die Evaluierenden eine Herausforderung dar. Klassische Evaluationsansätze gehen von klar definierten Evaluationsgegenständen und relativ stabilen Bedingungen aus; sie sind daher nur bedingt brauchbar für die Evaluation von innovativen, sich ständig anpassenden Maßnahmen unter veränderlichen Rahmenbedingungen. Für diese Fälle wurde von Michael Patton unter Einbezug von Konzepten aus der Komplexitätsforschung und der Systemtheorie die Developmental Evaluation (oder in der deutschen Übersetzung ‚evolutive Evaluation‘) entwickelt. Die breitere Erprobung dieses Ansatzes steht noch am Anfang (Lam/Shulha 2015).

Der folgende Beitrag geht kurz auf die Entstehung der Developmental Evaluation ein und nimmt eine Begriffsbestimmung in Abgrenzung von anderen Ansätzen vor. Er beschreibt zentrale Prinzipien und Konzepte der Developmental Evaluation und gibt einen Überblick über Anwendungsgebiete. Diese Ausführungen basieren stark auf dem grundlegenden Werk zu Developmental Evaluation von Patton (2010), beziehen sich aber auch auf andere Evaluationsforscherinnen und -forscher, die mit diesem Ansatz gearbeitet haben.

Im Anschluss wird die Anwendung von Developmental Evaluation anhand eines Fallbeispiels aus der Praxis der Autorin beschrieben. Es wird erörtert, warum dieser Evaluationsansatz in diesem Fall besonders geeignet erschien und wie vorgegangen wurde. Insbesondere werden auch Rolle und Aufgaben der Evaluatorin/des Evaluators, methodische Vorgangsweisen und organisatorische Rahmenbedingungen vorgestellt. Schließlich werden die Erfahrungen reflektiert und es wird ein Fazit zu Chancen und Grenzen von Developmental Evaluation, insbesondere in den durch das Fallbeispiel abgedeckten Anwendungsgebieten gezogen. Damit möchte der Bei-

Nutzen einer wirkungsorientierten Zielsystematik

Bindeglied zwischen Evaluation und Programmsteuerung

Stefan Schmidt

Dieser Artikel richtet sich in erster Linie an Evaluatorinnen und Evaluatoren, aber auch an Leitungen von Förderprogrammen bzw. Koordinatoren und Koordinatorinnen von Netzwerken. Die vorgestellte wirkungsorientierte Zielsystematik kann sowohl für die Evaluation als auch für die Steuerung von Programmen hilfreich sein. Sie wurde entwickelt, um die Verbindung von richtungsweisenden Leitbildern und allgemeinen Leitzielen mit konkreten Teilprojekten herzustellen und die einzelnen Teilprojekte von Förderprogrammen kohärent in das Gesamtkonzept einzufügen. Die wirkungsorientierte Zielsystematik schafft die Voraussetzungen dafür, prozessbegleitend und bilanzierend Daten zur Zielerreichung für Projekte und Maßnahmen zu erheben, um nachvollziehen zu können, ob sie die auf Wirkungen fokussierten Ziele erreicht. Die für die Entwicklung eines wirkungsorientierten Zielsystems notwendigen Klärungsprozesse können zudem eine konstruktive Zusammenarbeit der beteiligten Akteure fördern. Deshalb ist die wirkungsorientierte Zielsystematik besonders für Netzwerke mit einem heterogenen Akteursfeld geeignet. Aus Gründen der Lesefreundlichkeit wird in diesem Artikel

- der Begriff ‚Programm‘ verwendet, wenn es sich um ein intentional aufeinander bezogenes Bündel von Aktivitäten, Interventionen, Maßnahmen, Projekten oder Teilprogrammen handelt.
- der Begriff ‚Teilprojekt‘ verwendet, wenn es um örtlich und/oder zeitlich begrenzte Interventionen, Projekte und Maßnahmen geht, die Teil eines übergeordneten Programms sind.

Erfolgreich angewandt wurde die Vorgehensweise unter anderem im Bildungsnetzwerk ‚Lernen Erleben In Freiburg – LEIF‘. Wichtige Aspekte werden in diesem Artikel exemplarisch anhand der Erfahrungen in dieser Bildungslandschaft erklärt. Bildungslandschaften, in manchen Regionen wird auch der Begriff Bildungsnetzwerk verwendet, verfolgen das Ziel, Chancengerechtigkeit im Bildungswesen sicherstellen. Um dies zu erreichen, wird derzeit deutschlandweit in vielen Regionen ein systematisches Bildungsmanagement mit effektiven und effizienten Vernetzungsstrukturen zwischen bildungsrelevanten Institutionen aufgebaut. Dabei sind Kinder und Jugendliche in allen Bildungslandschaften eine wichtige Zielgruppe. Meist streben die Regionen aber perspektivisch an, das Lernen im gesamten Lebensverlauf in den Blick zu nehmen, so wie dies u. a. in der Stadt Freiburg im Rahmen einer Bundesprogrammförderung* erfolgte. Beteiligt sind neben Schulen in der Regel die Bildungseinrichtungen an den Übergängen (Kindertagesstätten, Ausbildungsbetriebe, Fachschulen, Hochschulen), darüber hinaus oft auch Einrichtungen der

Der Umgang mit Monitoring in der Evaluation

Katharina Klockgether

Unter Monitoring verstehen wir die wiederholte Erfassung von Daten zu gleichen Merkmalen eines Programms, in der Regel dessen Outputs, zum Zweck der Steuerung und Rechenschaftslegung. Obwohl es sich dabei um eine genuine Aufgabe von Politik- bzw. Programmverantwortlichen handelt, spielt das Monitoring in verschiedener Hinsicht eine bedeutende Rolle für die Evaluation von Programmen: Monitoringdaten unterstützen die Gegenstandsbestimmung und werden zur Beantwortung von Evaluationsfragestellungen herangezogen. Nicht selten werden zudem das Monitoring selbst bzw. eine Beratung bzgl. eines angemessenen Monitorings zur Aufgabe der Evaluation gemacht. In diesem Artikel wird die Entwicklung von Monitoring-Instrumenten im Rahmen von Evaluationen beschrieben und auf Monitorings als Datenquelle eingegangen. Dazu werden, nach einer grundsätzlichen Begriffsklärung und Charakterisierung, zunächst zwei Beispiele von Monitoring-Instrumenten präsentiert (Output- und Bedingungsmonitoring), die in Evaluationen durch Univation Anwendung fanden. Daraufhin werden exemplarisch die Entwicklung, Implementierung und Begleitung eines Monitoring-Instruments in einer Programmevaluation mit den sich ergebenden Herausforderungen dargestellt. Danach werden Hinweise für Evaluierende in Bezug auf die Nutzung, Umsetzung und Planung von Monitorings gegeben. Abschließend werden einzelne Monitorings vorgestellt, die in Deutschland in den Bereichen Gesundheit, Arbeitsmarkt, Bildung und Landwirtschaft umgesetzt werden. Dabei ist das Ziel des Artikels, die Chancen und Risiken von Monitorings für Evaluierende aufzuzeigen und mit der Darstellung der Erfahrungen aus der Praxis bei eigenen Vorhaben zu unterstützen.

1. Monitoring und Evaluation: Was ist das?

In Deutschland werden in vielen Bereichen Daten in sogenannten Monitorings gesammelt. Die Kultusministerkonferenz betreibt bspw. seit 1997 das Bildungsmonitoring zur Feststellung der Ergebnisse des Bildungssystems im Schulbereich. Das Robert-Koch-Institut erfasst über das Gesundheitsmonitoring Gesundheitszustand und -verhalten von Kindern, Jugendlichen und Erwachsenen. Die Gesellschaft für innovative Beschäftigungsförderung mbH veröffentlicht vierteljährlich für Arbeitsmarktakeure praxisgerecht aufgearbeitete Daten der amtlichen Statistik in einem ‚Arbeitsmarkt-Monitoring‘ für Nordrhein-Westfalen.¹ Diese kurze Auflistung, die sich leicht fortsetzen ließe, macht deutlich, dass Monitoring-Daten für Governance in verschiedenen Politikfeldern an Bedeutung gewinnen. Dies lässt ein Potenzial für

¹ Diese und weitere Monitorings werden im letzten Teil des Artikels kurz vorgestellt.

Interviews mit Jugendlichen

Herausforderungen in der Planung und Umsetzung

Matthias Sperling

Der vorliegende Text widmet sich der praktischen Durchführung von Face-to-Face-Interviews mit Jugendlichen im Rahmen von Evaluationen. Er zeigt dabei auf, warum Interviews mit Jugendlichen im Vergleich zu Interviews mit Erwachsenen ein besonderer Stellenwert zukommt, welche Herausforderungen und Problemlagen mit der Befragung von Jugendlichen in Evaluationen verbunden sein können und wie diesen in der Evaluationspraxis begegnet werden kann. Hierbei fließen zum einen Überlegungen aus der qualitativen Sozialforschung sowie der Evaluationstheorie ein, zum anderen werden Erkenntnisse aus der Evaluationspraxis gebündelt und weitergegeben.

Für die Durchführung von Interviews mit Jugendlichen werden grundlegende Überlegungen und Erfahrungen zur Organisation, zu Interviewformen und zur Leitfadententwicklung, zur Gesprächsführung, zum Umgang mit wiederkehrenden Herausforderungen, zu Dokumentationsformen sowie zum Datenschutz aufgezeigt. Die Zugangs- und Kontaktmöglichkeiten zu ‚Jugendlichen‘, deren Interesse bzw. Motivation zur Interviewteilnahme, der Ablauf eines Interviews sowie die Kompetanzanforderungen an die Interviewenden werden thematisiert. Darüber hinaus werden Hinweise zur Nachbereitung von Interviews, etwa zur Transkription im Evaluationskontext, zur Nachbegleitung von Jugendlichen sowie zu den (Selbst-)Reflexionskompetenzen der Interviewenden gegeben.

Jugend – als Zeitraum/Lebensabschnitt – wird im Rahmen des vorliegenden Textes als Lebensphase zwischen dem zwölften und dem 26. Lebensjahr definiert, die durch Identitätsfindungs- und Autonomieentwicklungsprozesse geprägt ist, wobei sie stets im individuellen Spannungsfeld zwischen Zukunfts- und Gegenwartorientierung liegt (Winkler 2005, S. 41 ff.).

1. Allgemeine Herausforderungen qualitativer Interviews

Eine erste Herausforderung besteht darin, ein Verständnis der Bedeutungszuschreibungen, die durch die Befragten im Einzelnen und in ihrer Gesamtheit getroffen wurden, zu entwickeln. Dazu müssen die Interviewenden die geäußerten Sinnzusammenhänge entschlüsseln (Was meint das Gegenüber?), die Zuschreibung dargestellter Bedeutungsträger (z. B. Wörter, Sätze, Vorgänge) identifizieren und sich um ein situatives Verständnis bemühen (In welchen Kontext ist das Geäußerte eingebettet? In welchem Rahmen findet das Interview statt?). Die Erschließung von Bedeutungszuschreibungen ist eine komplexe Aufgabe, da die Interviewenden nicht nur die individuellen Vorstellungen und Meinungen des Gegenübers, sondern auch

Mit Gruppendiskussionen kollektive Orientierungen rekonstruieren

Ute B. Schröder

Gruppendiskussionen liegen im Trend. In den letzten Jahren hat sich das Interesse sowohl im sozialwissenschaftlichen Bereich als auch bei Evaluationen verstärkt, Gruppendiskussionen als Erhebungsinstrument einzusetzen. Die Frage, die damit einhergeht ist, welches Wissen mit Gruppendiskussionen erhoben werden kann. Ist es die Summation von Einzelwissen, das (kosten- und zeitgünstig) in einer Gruppe erhoben wird? Sind es Theorien und Meinungen bspw. über Evaluationsgegenstände? Oder wird im Kontext einer Gruppendiskussion noch ein ganz anderes Wissen reproduziert, das in der Diskussion der Teilnehmenden zum Ausdruck kommt?

Im folgenden Beitrag steht diese weitere Art an Wissen, das so genannte implizite Wissen und die darauf basierenden kollektiven Orientierungen von Gruppen und dessen Bedeutung im Rahmen von Evaluationen, im Fokus. Es wird aufgezeigt, dass in Gruppendiskussionen gemeinsam geteilte Erfahrungen, so genannte kollektive Orientierungsmuster aktualisiert werden. Diese sind nicht nur an ein tatsächliches gemeinsames Handeln gebunden, sondern werden auch von Menschen geteilt, die bestimmte Erfahrungen gemeinsam haben, ohne im direkten Kontakt miteinander zu stehen. Jene Orientierungen bzw. jenes Wissen ist den Beteiligten nicht unbedingt reflexiv zugänglich, sondern vielmehr implizit. Es strukturiert jedoch die tägliche Handlungspraxis.

Anhand von Beispielen aus der Evaluationspraxis wird verdeutlicht, inwieweit die Rekonstruktion von kollektiven Orientierungen und impliziten Wissens für Evaluationen von Bedeutung sein kann. In diesem Zusammenhang werden Fragen bezüglich der Ermittlung von Wirkungen und Fragen der Bewertung von Evaluationsgegenständen thematisiert. Die Frage, die dahinter steht ist, inwieweit die Ebene des impliziten Wissens stärker für Evaluationen genutzt werden kann.

1. Die Gruppendiskussion – eine Methode der Datengewinnung aus der Sozialforschung

Gruppendiskussionen sind eine Methode der empirischen Sozialforschung, bei der – im Gegensatz zu Erhebungen mit einzelnen Personen – Meinungen, Theorien, Wissen und Erfahrungen einer Gruppe erhoben bzw. die Kommunikation in der Gruppe untersucht werden können. Bei der Zusammenstellung einer Gruppe für eine Gruppendiskussion wird zwischen zwei Arten unterschieden: der Realgruppe und der künstlichen Gruppe.

Die Anwendung der Grounded Theory in Evaluationen

Marianne Lück-Filsinger

Evaluation ist in einer allgemeinen Bestimmung als eine professionelle Dienstleistung zu verstehen, die auf der Grundlage empirischen Wissens jeweils zu bestimmende Gegenstände wie Politiken, Programme, Projekte, Maßnahmen, institutionelle Praxen u. a. systematisch und transparent beschreibt, analysiert und bewertet, um Entscheidungen zur ermöglichen (Widmer/Beywl/Fabian 2009; Stockmann/Meyer 2010). Die Zentralität empirischen Wissens verweist auf eine deutliche Nähe von Evaluation zur Sozialforschung, wenngleich die Schwierigkeiten zu beachten sind, empirisch gestützte Bewertungen mit intersubjektivem Geltungsanspruch zu gewinnen (Kromrey 2008). Bohnsack (2006) begreift Evaluation als einen eigenständigen Typ sozialwissenschaftlicher Forschung, schlägt aber mit Blick auf die Bewertungsproblematik eine Erweiterung der Verfahren der qualitativen Sozialforschung um eine Methodik der Moderation vor (Bohnsack/Nentwig-Gesemann 2010). Evaluation ist folglich eng mit der empirischen Sozialforschung in ihrer quantitativen und qualitativen Variante verknüpft oder anders formuliert, es geht um „Evaluation durch Forschung“ (Kromrey 2008).

Dieser Beitrag knüpft an die qualitative Sozialforschung an und ist fokussiert auf die qualitative Analyse empirischen Materials mit dem Verfahren der Grounded Theory.

Die qualitative Sozialforschung kennt eine Fülle von Forschungsstilen bzw. Auswertungsverfahren (für einen Überblick vgl. etwa Flick 2011; Bohnsack 2014). Im Evaluationskontext ist die Inhaltsanalyse (Mayring 2003) am weitesten verbreitet. In den letzten Jahren haben Ralf Bohnsack und seine Kolleginnen und Kollegen die Dokumentarische Methode der Interpretation im Zusammenhang mit der Methode der Gruppendiskussion fruchtbar gemacht (vgl. Bohnsack/Nentwig-Gesemann 2010; Lamprecht 2012 oder auch den Beitrag von Schröder in diesem Band). Susanne Giel hat jüngst Hinweise zu den Nutzungspotenzialen der Grounded Theory gegeben (2013, S. 200 ff.).

Im Folgenden soll vor dem Hintergrund eigener Evaluationserfahrungen die Produktivität dieses Forschungsstils und seiner Kodiervverfahren für das qualitativ-methodische Vorgehen im Evaluationszusammenhang gezeigt werden. Im Übrigen betonen Strauss/Corbin (1996) hinsichtlich der praktischen Relevanz dieses Forschungsstils selbst, dass es in der Grounded Theory nicht nur um die „Erklärung von Wirklichkeit“ geht, sondern dass sich diese darüber hinaus auch „zur Handlungskontrolle eignet“ (ebd., S. 8). Die Ausführungen basieren auf eigenen Erfahrungen

Evaluation und das Web

Elitsa Uzunova

Die Entwicklung von Web-Technologien ist rasant und passiert teilweise unbemerkt. Dies wirkt sich nicht nur auf unser Privatleben aus, sondern auch auf die Art und Weise, wie Menschen arbeiten. Die neuen Technologien sind intuitiv und effizient, und eröffnen ihrerseits zahlreiche Möglichkeiten zur Entwicklung neuer und zur Überarbeitung bestehender Geschäftsprozesse. Ihre rechtzeitige Anwendung kann den Wettbewerb in ganzen Wirtschaftsbereichen schnell verändern. Neue technologische Möglichkeiten können auch bei der Durchführung von Evaluationen dienlich sein und den Evaluationsprozess in sinnvoller Weise erweitern. Bislang ist ihre Auswirkung am deutlichsten im Bereich der Datenerhebung und -auswertung spürbar, auch wenn heutzutage noch nicht die ganze Palette von möglichen Technologien ausgenutzt wird. Fragen Sie sich jedoch: Wann haben Sie zum letzten Mal einen Fragebogen per Post zugeschickt? Wie haben Sie die letzte Fokusgruppe dokumentiert? Mit einem Papier-und-Bleistift-Protokoll? Die neuen Entwicklungen bringen selbstverständlich nicht nur Vorteile, sondern auch neue Herausforderungen mit sich. Im Bereich Evaluation (und nicht nur) stellt der adäquate Schutz personenbezogener Daten (vgl. Jelitto in diesem Band) vielleicht die größte Herausforderung dar. Eine andere Herausforderung (und gleichzeitig Grund für eine schnellere Einführung von Web-Technologien) stellt die sich verändernde Welt der zu evaluierenden Programmen (oder Organisationen) dar – sie sind nämlich schon online. Sie verbreiten ihre Erfolge als Videos im Netz, sie bieten Online-Kurse an, oder haben sogar ihre eigene App. In diesem Artikel wird ein Überblick über verschiedene Tools angeboten, die im Bereich Evaluation zum Einsatz kommen und den Nutzen für alle an einer Evaluation Beteiligten erhöhen können. Durch das Internet haben sich Prozesse der Datenerhebung und Datenauswertung in den letzten Jahren spürbar verändert. Das Web bietet neue Möglichkeiten zur Datengenerierung und bestehende Methoden und Vorgehensweisen lassen sich für das Web optimieren. Evaluierenden sollen diese Möglichkeiten in diesem Artikel nähergebracht werden. Der Text ist nach Anwendungsbereichen gegliedert – zunächst werden webtechnologische Möglichkeiten im Bereich Datenerhebung behandelt, dann in den Bereichen der Auswertung und der Ergebnispräsentation. Es werden jeweils die Vor- und Nachteile, die mit jeder technologischen Lösung verbunden sind, diskutiert. Auch werden Technologien kommentiert, die nicht direkt im Evaluationsprozess eingesetzt werden, sondern für den Wissenstransfer in der Evaluationscommunity förderlich sein können.

Bewertungsverfahren in der Evaluation

Vorgehensweisen und Praxisbeispiele

Susanne Mäder

1 Bewerten, aber wie?

Einer der zentralen Unterschiede zwischen Sozialforschung und Evaluation wird häufig darin gesehen, dass Evaluationen angemessene Bewertungen der beobachteten Praxis bereitstellen sollen. In seinem Beitrag „Qualitative Evaluationsforschung – was heißt hier Forschung?“ führt Lüders (2006, S. 51) aus, dass von Evaluationsforschung „eben nicht nur [die] ordentlichen empirischen Analysen und Rekonstruktionen sozialer Sachverhalte [erwartet werden], sondern eben auch eine besonders verlässliche Qualität der Bewertung dessen, was da beobachtet und analysiert worden ist“. Auch nach Haubrich u. a. (2005) unterscheidet sich Evaluation „von sonstigen Formen, meist als ‚wertfrei‘ verstandener sozialwissenschaftlicher Forschung insbesondere durch ihren Bewertungsanspruch“ (S. 1).

Schaut man sich die Verständnisweisen von Evaluation in der Literatur an, so weist die Mehrheit der Autorinnen und Autoren eine Bewertung der empirischen Daten als Aufgabe von Evaluationen aus (Shaw/Greene/Mark 2006; Widmer/Beywl/Fabian 2009; Widmer/de Rocchi 2012, Stockmann/Meyer 2010, Farrokhzad/Mäder 2014). Ebenso ist es in der Gesellschaft für Evaluation Konsens, dass unter Evaluation „die systematische Untersuchung des Nutzens oder Wertes eines Gegenstandes“ (DeGEval 2008, S. 15) verstanden wird. In ihren Standards für Evaluation (DeGEval 2008) wird eine Transparenz der Werte gefordert, auf deren Grundlage die Bewertungen vorgenommen werden. In den einleitenden Erläuterungen zu den Standards wird angeführt, dass eine professionelle Evaluation auf eine nachvollziehbare Bewertung ihres Gegenstands zielt, so dass dessen Güte und/oder Nutzen möglichst genau bestimmt werden kann (DeGEval 2008, S. 13).

In der Durchführung von Evaluationen ist es somit eine wichtige Aufgabe, zu Bewertungen zu gelangen. Auffällig ist jedoch angesichts der postulierten Wichtigkeit die eher stiefmütterliche Behandlung des Themas in der Evaluationsliteratur und in der Evaluationspraxis. In einschlägigen Werken zur Evaluation (wie bei Bortz/Döring 2006) wird die Frage, wie man zu Bewertungen gelangt, teilweise ausgeklammert. In der Praxis werden teils keine (expliziten) Bewertungen vorgenommen oder ihr Zustandekommen ist nicht nachvollziehbar. Demgemäß fordert Lüders bereits 2006 für die Evaluationsforschung einen methodologischen und methodischen Diskurs bezüglich der Frage, wie Bewertungen systematisch erfolgen können. Auch Kromrey (2007) wundert sich über das Stillschweigen innerhalb der Evaluationscommunity angesichts der aus seiner Sicht bestehenden Problematik einer empirisch gestützten Gewinnung von Wertaussagen, da die „unmittelbare Begründung