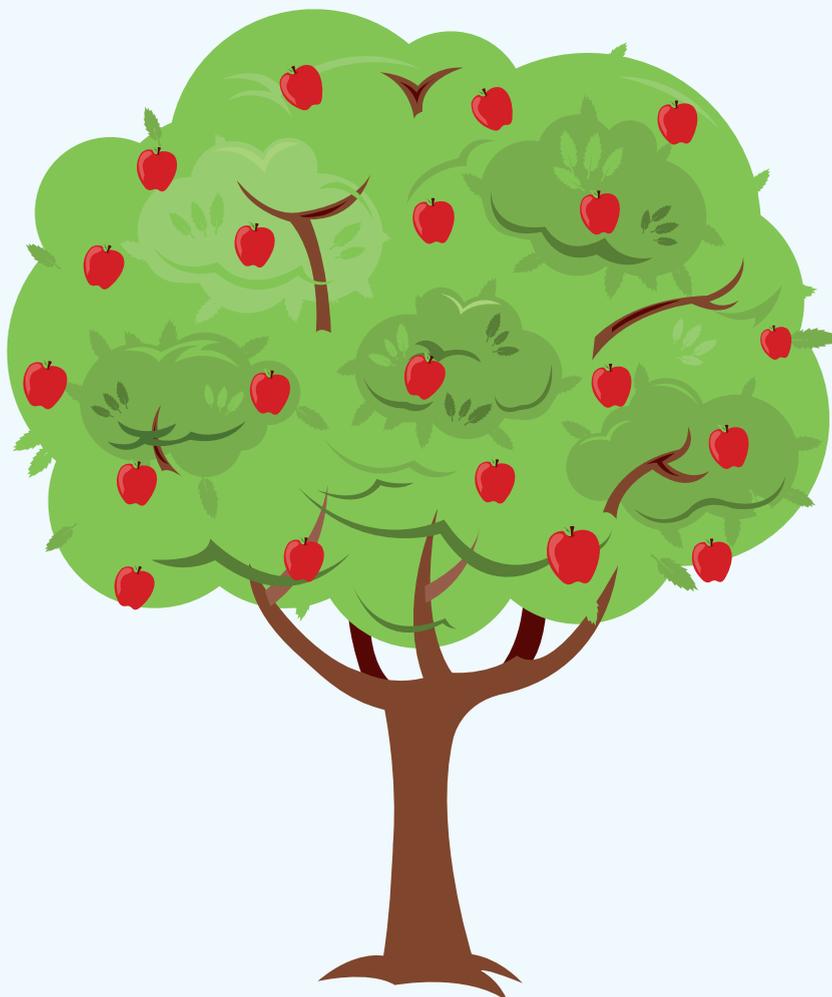


2. Auflage

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

Evaluationspraxis

Professionalisierung – Ansätze – Methoden



WAXMANN

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

Evaluationspraxis

Professionalisierung –
Ansätze –
Methoden



Waxmann 2016
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.

2. korrigierte und ergänzte Auflage

Print-ISBN 978-3-8309-3528-5

E-Book-ISBN 978-3-8309-8528-0

© Waxmann Verlag GmbH, Münster 2016

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.

Vorwort zur 2. Auflage

Mit der Herausgabe unseres Bandes zur Evaluationspraxis (2015) wollten wir eine Lücke in der deutschsprachigen Evaluationsliteratur schließen. Rückmeldungen im Rahmen von Rezensionen sowie der Buchvernissagen in Köln und Berlin zeigen: Die Verknüpfung von praktischen Erfahrungen mit einer theoretischen Reflexion wird als wichtig betrachtet:

„Es gibt zwar etliche Handreichungen und Leitfäden zur Durchführung von Evaluationen, doch es mangelt an jener Literaturkategorie, die Erkenntnisse und Erfahrungen aus der Evaluationspraxis vor einem fundierten theoretischen und forschungsmethodischen Hintergrund bündelt und problem- bzw. nutzungsorientiert an die im praktischen Evaluationsfeld stehenden Akteure adressiert.“ Prof. Dr. Helga Marburger (TU Berlin)

Uns freut es sehr, dass Leserinnen und Leser dem Band ein hohes Nutzungspotenzial zubilligen:

*„Was verbirgt sich hinter dem bescheidenen Titel „Evaluationspraxis“? In der Tat eine Schatztruhe und ein Handwerkskoffer zugleich: Was die Leserin findet, ist das geballte klassische wie innovative Wissen der Top-Evaluationsinstanz „Univation“ in komprimierter und anwender*innen-freundlicher Form.“* Prof. Dr. Lilo Schmitz (Hochschule Düsseldorf)

Bestätigt fühlen wir uns darüber hinaus, dass Leserinnen und Leser in den Beiträgen des Bandes einen Beitrag zur Professionalisierung ausmachen:

„Insgesamt betrachtet liegt ein Werk vor, das einen instruktiven Einblick in vielfältige Formen der Evaluationspraxis gibt, die in den Beiträgen vor dem Hintergrund von Evaluationsansätzen und methodologischen Entscheidungen systematisch reflektiert werden. [...] Der ohne Einschränkung lesenswerte Band leistet alles in allem einen sehr wertvollen Beitrag zur Profilbildung von Evaluation und sollte in den einschlägigen Diskursen in der Evaluationscommunity rezipiert und diskutiert werden.“ Jessica Prigge (HTW Saar)

Die vielen Leserinnen und Leser ermöglichen uns, eine zweite Auflage zu erstellen. Wir haben diese Gelegenheit genutzt, um Fehler zu korrigieren und einzelne Passagen inhaltlich zu aktualisieren. Wir möchten Wolfgang Beywl in diesem Zusammenhang für seine zahlreichen hilfreichen Hinweise danken!

Die Hervorhebung der Glossarbegriffe ist in der ersten Auflage leider grafisch nicht zufriedenstellend gelungen. Hier haben wir uns neben der Hervorhebung in grau zu einem zusätzlichen Fettdruck und einem Hinweis Pfeil entschieden. Wir hoffen, die Glossarbegriffe auf diese Weise mehr hervorzuheben, ohne den Lesefluss zu stark zu beeinträchtigen.

Neu hinzugekommen ist der Index im Anhang des Buches. Dort können zentrale Begriffe nachgeschlagen werden.

Köln, den 28.06.2016

Die Herausgeberinnen Susanne Giel, Katharina Klockgether und Susanne Mäder

Inhalt

Evaluationspraxis reflektieren – Einleitung und Überblick	9
<i>Susanne Giel/Katharina Klockgether/Susanne Mäder</i>	

Professionalisierungsinstrumente

Beitrag zu einer konsistenten Evaluationsterminologie – das Glossar von Univation	21
<i>Melanie Niestroj/Wolfgang Beywl/Berthold Schobert</i>	

Checklisten zur Planung und Steuerung von Evaluationen	31
<i>Katharina Klockgether</i>	

Informationssicherheit in der Evaluation	43
<i>Marc Jelitto</i>	

Interkulturelle Kompetenz – ein weiterer Baustein für das professionelle Handeln in Evaluationen	55
<i>Schahrzad Farrokhzad</i>	

Schwarm-Evaluation als Steuerungsansatz	73
<i>Dörte Schott und Wolfgang Beywl</i>	

Evaluationsansätze

Der Programmbaum als Evaluationsinstrument	89
<i>Samera Bartsch/Wolfgang Beywl/Melanie Niestroj</i>	

Wirkungen auf der Spur mit Programmtheorien	113
<i>Susanne Giel</i>	

Developmental Evaluation nach Michael Patton Begriffsbestimmung und Reflexion der praktischen Anwendung	133
<i>Maria Gutknecht-Gmeiner</i>	

Evaluationsmethoden

Nutzen einer wirkungsorientierten Zielsystematik Bindeglied zwischen Evaluation und Programmsteuerung	159
<i>Stefan Schmidt</i>	
Der Umgang mit Monitoring in der Evaluation	177
<i>Katharina Klockgether</i>	
Interviews mit Jugendlichen Herausforderungen in der Planung und Umsetzung	199
<i>Matthias Sperling</i>	
Mit Gruppendiskussionen kollektive Orientierungen rekonstruieren	219
<i>Ute B. Schröder</i>	
Die Anwendung der Grounded Theory in Evaluationen	237
<i>Marianne Lück-Filsinger</i>	
Evaluation und das Web	255
<i>Elitsa Uzunova</i>	
Bewertungsverfahren in der Evaluation Vorgehensweisen und Praxisbeispiele	279
<i>Susanne Mäder</i>	
Autorinnen und Autoren	299
Index	303



Die in diesem Band verwendeten Evaluationsfachbegriffe sind im Glossar der wirkungsorientierten Evaluation von Univation definiert. Die Begriffe sind → mit Pfeil, in grau und fett gekennzeichnet und können unter www.eval-wiki.org eingesehen werden.

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 → **Wirksamkeiten**, 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 Interpretation 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 *Schahrazad 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. Der Artikel 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 mit einem Pfeil markiert sowie fett und 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.

Literatur

- Balzer, Lars/Beywl, Wolfgang (2015): *evaluiert – Planungsbuch für Evaluationen im Bildungsbereich*. Bern: hep verlag.
- Beywl, Wolfgang (1988): *Zur Weiterentwicklung der Evaluationsmethodologie*. Frankfurt: Lang.
- Beywl, Wolfgang (2006): *Evaluationsmodelle und der Stellenwert qualitativer Methoden*. In: Flick, Uwe (Hrsg.): *Qualitative Evaluationsforschung. Konzepte – Methoden – Umsetzung*. Reinbek: Rowohlt, S. 92–116.
- Beywl, Wolfgang/Bestvater, Hanne/Friedrich, Verena (2011): *Selbstevaluation in der Lehre*. Münster: Waxmann.
- Beywl, Wolfgang/Kehr, Jochen/Mäder, Susanne/Niestroj, Melanie (2007): *Evaluation. Schritt für Schritt: Planung von Evaluationen*. Darmstadt: hiba.
- Beywl, Wolfgang/Niestroj, Melanie (2009): *Der Programmbaum. Landmarke wirkungsorientierter Evaluation*. In: Univation (Hrsg.): *Das A-B-C der wirkungsorientierten Evaluation*. 2. Aufl., Köln: Univation – Institut für Evaluation GmbH, S. 137–149.
- Beywl, Wolfgang/Schepp-Winter, Ellen (2000): *Materialien zur Qualitätsentwicklung, Zielgeführte Evaluation von Programmen, Materialien zur Qualitätssicherung in der Kinder- und Jugendhilfe*, hg. vom Bundesministerium für Familie, Senioren, Frauen und Jugend, Heft Nr. 29, Berlin. http://univation.org/download/QS_29.pdf [Stand: 24.06.2015]
- Beywl, Wolfgang/Speer, Sandra (2004): *Standards for Evaluation Practices, On the Way to Develop Standards for Program Evaluation in Vocational Education and Training Contexts*. In: Descy, Pascaline/Tessaring, Manfred (Hrsg.): *Third Research Report on Vocational Education and Training*. CEDEFOP, Luxembourg.
- Beywl, Wolfgang/Speer, Sandra/Kehr, Jochen (2003): *Ansätze zur Wirkungskontrolle im Rahmen der Armut- und Reichtumsberichterstattung, Bundesministerium für Gesundheit und Soziale Sicherung (BMGS) (Hrsg.)*. Bonn.
- Brand, Tasso (2009): *Evaluation in Deutschland. Professionalisierungsstand und -perspektiven*. Münster u. a.: Waxmann.
- DeGEval – Gesellschaft für Evaluation e. V. (2008a): *Standards für Evaluation*. http://www.degeval.de/fileadmin/user_upload/Sonstiges/STANDARDS_2008-12.pdf [Stand: 27.05.2015].
- DeGEval – Gesellschaft für Evaluation e. V. (2008b): *Empfehlungen für die Aus- und Weiterbildung in der Evaluation. Anforderungsprofile an Evaluatorinnen und Evaluatoren*. http://www.degeval.de/fileadmin/Publikationen/Publikationen_Homepage/DeGEval_-_Empfehlungen_Aus-_und_Weiterbildung.pdf [Stand: 21.05.2015]
- Farrokhzad, Schahrzad/Mäder, Susanne, (2014): *Nutzenorientierte Evaluation: Ein Leitfaden für die Arbeitsfelder Integration, Vielfalt und Toleranz*. Münster: Waxmann.
- Giel, Susanne (2013): *Theoriebasierte Evaluation. Konzepte und methodische Umsetzung*. Münster: Waxmann.
- Keller-Ebert, Cornelia/Kißler, Mechtilde/Schobert, Berthold (2005): *Evaluation praktisch! Wirkung überprüfen, Maßnahmen optimieren, Berichtsqualität verbessern*. Darmstadt: hiba.
- Kromrey, Helmut (2001): *Evaluation – ein vielschichtiges Konzept. Begriff und Methodik von Evaluierung und Evaluationsforschung. Empfehlungen für die Praxis*. In: *Sozialwissenschaften und Berufspraxis*. Bd. 24., S. 105–131.

- Mäder, Susanne (2013): Die Gruppendiskussion als Evaluationsmethode – Entwicklungsgeschichte, Potenziale und Formen. In: Zeitschrift für Evaluation. 12. Jg., Heft 1, Münster: Waxmann. S. 23–51.
- Patton, Michael Q. (1997): Utilization-focused Evaluation. Thousand Oaks (Calif.): Sage.
- Stake, Robert E. (1995): The art of case study research. Thousand Oaks (Calif.): Sage.
- Widmer, Thomas/Beywl, Wolfgang (2000): Die Übertragbarkeit der Evaluationsstandards auf unterschiedliche Anwendungsfelder. In: Joint Committee on Standards for Educational Evaluation (Hrsg.): Handbuch der Evaluationsstandards. Opladen: Leske und Budrich, S. 243–257.

Professionalisierungsinstrumente

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 427 evaluationsrelevante Begriffe definiert sind und das allen Interessierten kostenfrei zur Verfügung steht.

2. Zielsetzung des Beitrags

Das Team von Univation besteht von Anfang an aus Personen mit sehr unterschiedlichen professionellen Hintergründen. Mit dem Wachsen des Teams nimmt diese professionelle Heterogenität zu. Darüber hinaus arbeiten wir intensiv mit zahlreichen Partnerinnen und Partnern zusammen, die als externe Fachpersonen vertiefte Kenntnisse aus → **Feldern** mit eignen Fachsprachen, jedoch nicht notwendig eigene Evaluationsexpertise einbringen. So befruchtend eine solche Zusammenarbeit in vielerlei Hinsicht ist, treten dabei immer wieder Herausforderungen gerade auch auf der Ebene der sprachlichen Verständigung auf.

Univation verfolgt einen dezidiert an der Erzeugung von → **Nutzen** orientierten → **Evaluationsansatz** und setzt dabei unterschiedliche → **Evaluationsmodelle** sowie verschiedenste → **Methoden** und Instrumente u. a. zur Steuerung von Evaluationen, zur → **Datenerhebung** oder → **-auswertung** ein. (In den Beiträgen dieses Bandes finden sich detaillierte Informationen dazu.) Diese Grundlagen müssen innerhalb immer wieder ergänzter und wechselnder Teams sowie gegenüber externen Mitarbeitenden vermittelt werden. Besonders in unseren Evaluationsangeboten, in → **Evaluationsberichten**, aber auch in internen Prozessbeschreibungen, Checklisten (vgl. Klockgether zu Checklisten in diesem Band) und anderen Publikationen ist ein einheitlicher Wortgebrauch sehr wichtig, um mit den jeweiligen Adressaten klar und konsistent zu kommunizieren. Wir haben die Erfahrung gemacht, dass ein einheitlich genutztes und verbindliches Vokabular dabei eine wertvolle Basis ist.

Da es Univation ein Anliegen ist, zur Weiterentwicklung von Theorie und Methodik der Evaluation beizutragen, ermöglichen wir allen interessierten Personen die Nutzung unseres ‚Wort-Schatzes‘ (fachsprachlich: Thesaurus), was dazu beigetragen hat, dass die Terminologie in den letzten Jahren auch über unser Team hinaus zunehmend Anwendung findet. Dieser Beitrag zeichnet die Entwicklung des heutigen ‚Eval-Wiki: Glossar der Evaluation‘ nach und zeigt auf, welchen Niederschlag die Terminologie in unserer Arbeit und unseren Produkten findet, wie sie in der deutschsprachigen Evaluationscommunity genutzt wird, welchen Beitrag sie zur weiteren Professionalisierung leisten kann und welche Mitgestaltungsmöglichkeiten Mitglieder der Community wahrnehmen können.

3. Entstehung des Eval-Wiki: Glossar der Evaluation

Die Initiative der bis heute intensiv betriebenen Auseinandersetzung mit Evaluationsterminologie bei Univation ging von Wolfgang Beywl, dem wissenschaftlichen Leiter des Instituts, aus. Bereits im Zuge seiner Dissertation Ende der 1980er Jahre beschäftigte er sich mit der US-amerikanischen Evaluationstheorie und -geschichte (Beywl 1988). Dabei musste er die englischen Originaltexte in Teilen übersetzen und für Fachtermini angemessene deutsche Entsprechungen finden – zum damaligen Zeitpunkt gab es kaum originäre deutschsprachige Fachliteratur zur Evaluation. In

der Vorbereitung von Seminaren, die Ende der 1990er Jahre in Deutschland mit Michael Q. Patton stattfanden, wurden Begriffe mit Erläuterungen aus Pattons damals dritter Auflage von *Utilization-Focused Evaluation* (1997) als Vorbereitungsmaterial für die Teilnehmenden übersetzt. Dies setzte sich bspw. fort in der Übersetzung der → *Evaluationsstandards* des US-amerikanischen Joint Committee on Standards for Educational Evaluation (gemeinsam mit Thomas Widmer; JCSEE/Sanders 2000), deren deutschsprachige Buchpublikation bereits ein ausführliches Glossar enthält. Aus der kontinuierlichen Auseinandersetzung mit der englischsprachigen Evaluationsliteratur und der Wolfgang Beywl eigenen Leidenschaft für Sprache entstanden dabei auch neue deutsche bzw. eingedeutschte Fachbegriffe, die sich bis heute in unserem Sprachgebrauch gehalten haben (z. B. der → ‚*Kriterienpunkt*‘ als eine Übersetzung für das englische ‚*standard*‘, welches viele sehr unterschiedliche, in den deutschen Übersetzungen dann aufgelöste Bedeutungen hat, oder → ‚*Incomes*‘ als Komplement zu → ‚*Outcomes*‘).

Neben der Nutzung in verschiedenen Fachpublikationen Wolfgang Beywls erfuhr die Terminologie eine erste für die deutsche Evaluationscommunity relevante Anwendung bei der Erarbeitung der Standards für Evaluation der DeGEval, die bis 2001 koordiniert durch Wolfgang Beywl erfolgte. Die Ergebnisse des partizipativ angelegten Prozesses erschienen u. a. als Printversion (DeGEval 2002). In den Fußnoten des Kapitels ‚*Begleitmaterialien zu den Standards für Evaluation*‘ finden sich darin Anmerkungen zu den gewählten Begriffen, wie bspw. → ‚*Zwecke der Evaluation*‘ vs. → ‚*Ziele des Programms*‘ (ebd., S. 13).

Mit dem Team von Univation wurde die Terminologiarbeit fortgeführt. Anfang 2004 erschien im Bericht zum Forschungsprojekt ‚*Wirkungsorientierte Evaluation im Rahmen der Armut- und Reichtumsberichterstattung*‘ (Beywl/Speer/Kehr 2004), der u. a. verschiedene Evaluationsmodelle vorstellt und ihre Nutzbarkeit für Evaluationen in der Armut- und Reichtumsberichterstattung diskutiert, ein rund 50 Einträge umfassendes Glossar. Auf dessen Basis gab Univation Ende 2004 eine erste, auf rund 160 Einträge erweiterte Printversion des ‚*Glossars – Deutsch/Englisch – der wirkungsorientierten Evaluation*‘ im Eigenverlag heraus (Univation 2004). Nicht nur wurden dabei die Begriffsdefinitionen um den Verweis auf die englischen Fachbegriffe erweitert, auch wurde im Vorwort erstmals der Anspruch formuliert, ein „konsistentes begriffliches Instrumentarium“ zu entwickeln, das zentrale Begriffe der Evaluation sowie wichtige Begriffe, die Programme als unsere Gegenstände betreffen, aufgreift und die vielfältigen inhaltlichen Verknüpfungen der Begriffe durch Querverweise aufzeigt. Der Anspruch dabei war und ist, ein kontrolliertes Vokabular zu erstellen. Das beinhaltet u. a. die Anforderung, dass die Glossarbegriffe selbst, werden sie in den Definitionen anderer Glossarbegriffe genutzt, in der im Glossar definierten Bedeutung genutzt werden müssen. Andernfalls muss der betreffende Begriff durch ein Alltagssprachliches Synonym ersetzt werden. Gleichzeitig wurde nach außen hin betont, in den Begriffsdefinitionen keine Allgemeingültigkeit beanspruchen zu wollen, da sich zu manchen der enthaltenen Begriffe in der Literatur

Differenzen bis hin zu direkten Widersprüchlichkeiten finden.¹ Schon damals wurde eine erste – einfache – Online-Version des Glossars über die Univations-Homepage kostenfrei zur Verfügung gestellt, Mitglieder der Evaluationscommunity wurden um Ergänzungshinweise und Anmerkungen gebeten.

Zu dieser ersten Glossarpublikation erhielten wir von Personen aus der deutschsprachigen Evaluationscommunity sehr positive und zur Weiterarbeit ermunternde Rückmeldungen. Unter anderem stellten sie das Potenzial für die Weiterentwicklung der – insbesondere in Deutschland – immer noch jungen Profession Evaluation bspw. durch die Nutzung des Glossars in der Aus- und Weiterbildung von Evaluierenden heraus. Auch die damaligen Präsidenten der Rechnungshöfe im deutschsprachigen Raum, denen wir ein Exemplar der Broschüre zuschickten, reagierten positiv.

Eine auf 370 Begriffe erweiterte Printversion des Glossars erschien im Jahr 2009 (Univation 2009). Dabei flossen Erfahrungen aus weiteren Jahren der Evaluationsforschung und -praxis, der Weiterbildung von Evaluatorinnen und Evaluatoren² sowie intensive Auseinandersetzungen über einzelne Begriffe mit Fachkolleginnen und -kollegen ein. Auch wurden den Definitionen vieler Begriffe erstmals Literaturverweise zur tiefergehenden Auseinandersetzung mit der jeweiligen Thematik angefügt. Im Vorwort dieser 2. Auflage wurden der Entstehungshintergrund und der bereits genannte Anspruch des Glossars sowie die Grenzen seines Begriffsraumes detaillierter erläutert und es wurde auf weitere Glossare verwiesen.³ Zudem wurde im Rahmen eines abschließenden Artikels der Programmbaum als ein zentrales Instrument unserer Arbeit, in dem sich viele zentrale Begriffe im Zusammenhang wiederfinden, vorgestellt (vgl. Bartsch/Beywl/Niestroj in diesem Band). Das Printglossar 2009 wurde in einer Auflage von 1.000 Exemplaren gedruckt. Es konnte bei Univation direkt bestellt werden, wurde durch das Team jedoch auch bei vielen (Weiterbildungs-)Veranstaltungen und bei persönlichen Kontakten mit Evaluierenden, → **Auftraggebenden** und weiteren an Evaluation interessierten Personen verteilt.

Uns war zu diesem Zeitpunkt klar, dass es eine weitere Printversion aufgrund des hohen damit verbundenen Aufwandes und der fehlenden Aktualisierungsmöglichkeiten nicht geben würde. Mit der Wiki-Technologie ergab sich aus unserer Sicht eine Möglichkeit, das Glossar in ein Online-Angebot zu verwandeln, das eine hohe Nutzerfreundlichkeit und breite Verfügbarkeit bietet sowie optimale Bearbeitungs- und Weiterentwicklungsmöglichkeiten eröffnet. Im September 2010 wurde der interessierten Öffentlichkeit das ‚Eval-Wiki: Glossar der Evaluation‘ (<http://www.eval-wiki.org/glossar>) präsentiert. Als Neuerung fanden sich diesmal neben

1 Bspw. im Falle der Begriffe → ‚Outputs‘, ‚Outcomes‘, → ‚Impacts‘, die bspw. in Lehrbüchern der Schul- (bspw. Strahm 2008, S. 20) bzw. Politikevaluation (bspw. Thom/Ritz 2008, S. 221 f.) sehr unterschiedlich verwendet werden.

2 U. a. im Rahmen der Weiterbildungsstudiengänge ‚Evaluation‘ des Zentrums für universitäre Weiterbildung der Universität Bern.

3 Vgl. PDF der Einleitung unter http://www.univation.org/download/Univation_glossar_leseprobe.pdf [Stand: 19.02.2015]

den englischen (bedingt durch Wolfgang Beywls Tätigkeit in der Schweiz) auch die französischen Entsprechungen der Fachbegriffe. Die ergänzenden Literaturangaben wurden deutlich erweitert. Über die im Glossar der Evaluation gegebenen Diskussionsmöglichkeiten der Begriffe hinaus wollten wir insbesondere die Entwicklung neuer Begriffsdefinitionen für Fachpersonen über unseren Kreis hinaus öffnen. Mit der ‚Eval-Wiki: Glossar-Werkstatt‘ (<http://eval-wiki.org/werkstatt>) wurde dazu ein zweites Wiki angelegt, in dem Nutzende Begriffe vorschlagen und an Begriffsdefinitionen arbeiten können. Fertig definierte Begriffe daraus werden nach einer Endredaktion (v. a. um den Anspruch des konsistenten begrifflichen Instrumentariums zu schützen) in das Eval-Wiki: Glossar der Evaluation übertragen. Aktuell umfasst das Glossar der Evaluation 427 Begriffe, in der Glossar-Werkstatt finden sich 106 Einträge (mit unterschiedlichem Bearbeitungsgrad).

Die Begriffsdefinitionen im Hauptglossar sind relativ stabil, aber nicht unveränderlich. Mit zeitlichem Abstand gelingt es oft besser, den gemeinten Sachverhalt klar zu stellen, was zu leichten Anpassungen führt. Teils fallen aber auch Brüche bis hin zu missverständlichen Formulierungen auf, die korrigiert werden. Insofern spiegelt das Glossar auch unseren evaluationsfachlichen Lernprozess wider. Der Zeitpunkt der letzten Bearbeitung jeder Begriffsdefinition ist auf der jeweiligen Seite festgehalten und sollte beim Zitieren von Glossardefinitionen mit genannt werden.

Als jüngster Spin-off des Eval-Wiki entstand durch Wolfgang Beywl und Mitarbeitende seiner Professur an der Pädagogischen Hochschule FHNW ein Wiki als Begleitmedium zu den Veröffentlichungen von John Hattie (z. B. 2013). Es ist erreichbar über das Portal www.lernensichtbarmachen.net. Kern sind die Kurzfassungen der ca. 150 Faktoren, die nach Hattie in unterschiedlich starkem Ausmaß Einfluss auf Lernleistungen (von Schülerinnen und Schülern) nehmen. Darüber hinaus enthält dieses Wiki ca. 220 Glossarbegriffe, die – auch für fachliche Laien verständlich – Fachterme der Bildungs- und Lernforschung erläutern. Da John Hatties Kernthese ist, dass Lehrpersonen dann besonders wirkfähigen Unterricht machen, wenn sie dessen Wirkungen auf das Lernen der Schülerinnen und Schüler ‚evaluieren‘, gibt es zahlreiche Querverbindungen zwischen den beiden Glossaren.⁴

2015 wurden dem Glossar der Evaluation im Eval-Wiki weitere Ressourcen zur Seite gestellt. Es finden sich hier nun neben deutschsprachigen Evaluations-Checklisten (vgl. Klockgether in diesem Band) u. a. Hinweise zum Instrument Programmbaum, zu Evaluationsstandards sowie der Beginn einer weiterreichenden Quellensammlung zu Informations- und Selbstlernmöglichkeiten zur Evaluation

4 Für die Geschichtsschreibung sei auf eine gemeinsame Quelle verwiesen: John Hatties Frau Janet Clinton ist Programmevaluatorin, hat vielfach mit ihm gemeinsam (z. B. 2012) und eigenständig (z. B. 2014) veröffentlicht und sein Denken zu Evaluation offensichtlich stark beeinflusst. Einer ihrer akademischen Lehrer ist Michael Scriven, der zugleich geniale wie bei seinen Wortschöpfungen offensichtlich unerschöpfliche Mitbegründer der Evaluation als ‚Transdisziplin‘. Sein Evaluationsthesaurus (Scriven 1991), und die Key-Evaluation Checklist (Scriven 2007) sind wichtige Quellen für unser Eval-Wiki: Glossar der Evaluation.

im Internet. Wir haben den Anspruch, das Eval-Wiki systematisch zu pflegen und auszuweiten, um Evaluierenden, Auftraggebenden und weiteren Interessierten eine umfassende Ressourcensammlung zur Evaluation zu bieten. (Alle Informationen und Materialien sind erreichbar über www.eval-wiki.org.)

4. Nutzung des Glossars

Univation selbst bemüht sich, in allen Arbeitsdokumenten und Veröffentlichungen – Evaluationsberichten, unserer Homepage sowie Fachveröffentlichungen wie die Artikel dieses Sammelbands – um die durchgängige Verwendung der Begriffe wie im Glossar definiert. Als eine Art von Produkten, in denen auf die durchgängige Anwendung der Glossarbegriffe geachtet wird, seien bspw. die deutschen Übersetzungen von Checklisten des ‚Evaluation Checklist Project‘ des Evaluation Center der Western Michigan University genannt (vgl. Klockgether zu Checklisten in diesem Band). Buchpublikationen, die von Mitgliedern des Univations-Teams (mit-) herausgegeben werden, enthalten zudem mitunter Auszüge des Glossars, basierend auf der jeweils aktuellen Online-Version (bspw. Farrokhzad/Mäder 2014; Balzer/Beywl 2015).

Es muss kaum erwähnt werden, dass die sprachliche Konsistenz angesichts der vielen verschiedenen jeweils beteiligten Autorinnen und Autoren nicht immer einfach zu realisieren ist, insbesondere wenn diese Produkte (auch) durch Personen erstellt werden, die nicht zum Kernteam gehören. Eine ‚Qualitätskontrolle‘ ist unerlässlich. Diese gelingt am besten, wenn das Online-Glossar beim Schreiben von Texten in einem eigenen Fenster geöffnet ist und es ähnlich wie ein Online-Wörterbuch bei Übersetzungsarbeiten genutzt wird. Beim Schreiben von Texten wie beim Lesen von Evaluationsfachliteratur stoßen wir jedoch selbst immer wieder auf neue begriffliche Herausforderungen, die zu Ergänzungen oder Veränderungen der Glossardefinitionen führen.

Im Folgenden soll skizziert werden, welche Hinweise uns dazu vorliegen, wie das Glossar über unser Team hinaus, also in der deutschsprachigen Evaluationscommunity, genutzt wird, und wie damit unser Anspruch, zur Professionalisierung der Evaluation beizutragen, bisher erreicht werden konnte.

Im Rahmen der Jahrestagung der DeGEval 2011 in Linz (unter dem Titel ‚Partizipation – dabei sein ist alles!?’) stellten wir in einer Session des AK Aus- und Weiterbildung in der Evaluation (heute AK Professionalisierung) erste Nutzungsdaten zum Eval-Wiki: Glossar der Evaluation vor (das damals ein Jahr online zur Verfügung stand).⁵ In den Monaten März bis August 2011 verbuchten wir ca. 720 Besuche monatlich, darunter etwa 160 wiederkehrende Besuche von Personen v. a. aus Deutschland (ganz überwiegend), Österreich und der Schweiz. Die damalige Gesamtzahl von Abrufen der Startseite betrug rund 3.600. Heute (Mai 2016, also

5 Download der Präsentation unter http://www.degeval.de/images/stories/Jahrestagung/Linz_2011/Dokumentation/B1_Niestroj_Uzunova.pdf [Stand: 19.02.2015].

fünf Jahre später) wurde die Startseite des Eval-Wiki: Glossar der Evaluation gut 136.500 Mal aufgerufen (womit Glossar-Nutzende nicht erfasst werden, die direkt auf eine weiterführende Seite des Wikis, bspw. eine Begriffsdefinition zugreifen). In den letzten vier Monaten lag die Anzahl von monatlichen Besuchen im Mittel bei knapp 1.300, davon im Mittel knapp 180 wiederkehrende Besuche. (Bei allen Zahlen werden natürlich auch unsere eigenen Zugriffe erfasst.)

Zwischen Juli und September 2011 wurde außerdem eine kurze Online-Befragung zum Glossar unter Nutzenden durchgeführt (Aufruf und Link zum Online-Fragebogen wurden auf der Startseite des Eval-Wiki eingestellt), über die wir etwas über die Nutzenden und ihre Bewertung des Angebots herausfinden wollten. Es beteiligten sich 38 Personen, gut drei Viertel von ihnen Evaluierende und Personen aus Wissenschaft und Forschung, einzelne Studierende, eine Person aus der Gruppe der Auftraggebenden von Evaluation. Die meisten von ihnen besuchten das Glossar sporadisch (monatlich oder seltener). Ihre Interessen beim Besuch waren v. a. die Suche nach Klarheit über Begriffsdefinitionen (17 Mal genannt), die Nutzung von Verweisen für Lehre, Weiterbildungen etc. (5), die eigene Information/Weiterbildung (4), die Nutzung beim Schreiben von Texten (4) oder die Suche nach Übersetzungen der Begriffe (3). Die Stärken der Wikis lagen aus der Sicht der Antwortenden darin, Begriffsdefinitionen zu liefern/als Nachschlagewerk zu dienen (9 Nennungen), aktuell/schnell/übersichtlich, umfangreich bzw. verständlich zu sein (13), redigierte bzw. fachlich geprüfte Informationen zu liefern (3), kostenfrei zur Verfügung zu stehen (2) sowie eine Quelle für Übersetzungen, Literaturhinweise und inhaltliche Verlinkungen zwischen den Begriffen zu sein (jeweils 2). Lediglich eine Person sprach die Möglichkeit, sich an der Erarbeitung von Begriffsdefinitionen zu beteiligen, als eine Stärke des Angebots an. Verbesserungshinweise von einzelnen Antwortenden bezogen sich auf eine kritischere Diskussion der Begriffe bspw. über die Angabe von Alternativdefinitionen mit entsprechenden Quellen (welche in einzelnen Fällen im Online-Glossar angegeben sind, vgl. bspw. die ‚Variantentafel‘ zu verschiedenen Resultatsarten von Programmen unter http://www.eval-wiki.org/glossar/Resultate_%28eines_Programms%29). Mehr als die Hälfte der Antwortenden hatte das Online-Glossar schon weiterempfohlen.

Über dieses Blitzlicht hinaus erreichen uns aus der Evaluationscommunity informell immer wieder – zumeist wohlwollende und ermutigende – Rückmeldungen zum Glossar und Hinweise dazu, wo dieses eingesetzt wird (bspw. in der Lehre). Auch im Internet finden sich an verschiedenen Stellen von Seiten öffentlicher Institutionen Verweise auf unser Glossar, wie bspw. beim Bildungsserver Rheinland-Pfalz⁶ oder dem Niedersächsischen Bildungsserver (NiBiS)⁷, dem Schweizerischen Bundesamt für Gesundheit (BAG)⁸ oder dem Deutschen Bundesamt für Wirtschaft

6 Siehe Verweis unter <http://ines.bildung-rp.de/evaluation-und-qualitaetsentwicklung/ines-glossar.html> [Stand: 23.02.2015] ergänzend zu einzelnen eigenen Definitionen.

7 Siehe <http://portal.eval.nibis.de/nibis.php?menid=138> [Stand: 23.02.2015]

8 Siehe Verweis unter <http://www.bag.admin.ch/evaluation/02357/02603/index.html?lang=de> [Stand: 23.02.2015] ergänzend zu eigenen Definitionen.

und Ausfuhrkontrolle (BAFA)⁹. Auch einzelne Kolleginnen und Kollegen im Feld berufen sich auf das Eval-Wiki und verweisen darauf von ihren Webseiten.¹⁰

Schlussfolgern könnte man aufgrund der aufgeführten Hinweise, dass das Eval-Wiki: Glossar der Evaluation in der Community der Evaluierenden im deutschsprachigen Raum Beachtung findet und auf einem guten Weg ist, einen Beitrag zur Professionalisierung der Evaluation zu leisten. Allerdings haben wir mit dem Online-Glossar im Wiki-Format noch einen weiteren Anspruch verbunden: Die *partizipative* Weiterentwicklung der Evaluationsterminologie in der Form, dass sich Fachpersonen aktiv an den Begriffsdefinitionen beteiligen bzw. diese diskutieren. Dieser Anspruch konnte bis heute nicht eingelöst werden. In den seltenen Fällen, in denen sich Personen außerhalb des Univationsteams an der (Weiter-)Entwicklung einer Begriffsdefinition beteiligen, handelt es sich i. d. R. um externe Kolleginnen/Kollegen, mit denen wir gerade am betreffenden Thema arbeiten.

Es stellt sich in diesem Zusammenhang die Frage, ob es in der deutschsprachigen Evaluationscommunity den Wunsch gibt, einen Vorstoß zu einem einheitlichen Vokabular zu unternehmen. Deutlich wurde ein solcher während der partizipativen Erarbeitung der DeGEval-Standards für Evaluation 1999–2001 zumindest in der Gruppe von Expertinnen und Experten, die an der Entwicklung der Standards beteiligt waren. Von Seiten der DeGEval oder auch der Schweizerischen Evaluationsgesellschaft SEVAL, als den beiden Fachgesellschaften für den deutschsprachigen Raum, gibt es bis heute – über die bereits angesprochenen einzelnen Begriffsdefinitionen in den Anmerkungen zu den DeGEval-Standards hinaus – keine Bemühungen zur Vereinheitlichung der Terminologie. In den beiden fachlich sehr heterogen zusammengesetzten Gesellschaften wäre eine solche mit sehr langen und intensiven Klärungsprozessen verbunden. Jedoch könnte gerade hier eine Klärung die echte interdisziplinäre Verständigung befördern, die durch nebeneinander genutzte – und meist auch intern sehr wenig standardisierte – Fachsprachen der Disziplinen (u. a. Psychologie, Soziologie, Politikwissenschaft, Erziehungswissenschaft, Volks- und Betriebswirtschaftslehre, Verwaltungswissenschaften, Geographie) behindert wird.¹¹

Sollte es von ‚offizieller Seite‘ einen Vorstoß zur Begriffsklärung geben, könnte sich das Univations-Glossar als Ausgangspunkt anbieten. Solange verstehen wir es

9 Siehe http://www.bafa.de/bafa/de/weitere_aufgaben/evaluierung/links/eval_wiki_glossar_der_evaluation.html [Stand: 23.02.2015]

10 Siehe bspw. Evaluelab – Gesellschaft für empirische Sozialforschung mbH (<http://www.evaluelab.de/unsere-leistungen/evaluation>), WiDi-Kontor: Wissenschaftliche Dienstleistungen & Forschungsservice Dr. Eva Hampel (<http://www.widi-kontor.de/links.html>) oder Evaluationskultur Gabriela Dömötör (<http://www.evaluationskultur.ch/>).

11 Anfang 2015 informierte die DeGEval ihre Mitglieder darüber, dass zu der revidierten Fassung der Standards für Evaluation ein Glossar erstellt werden soll. Zwischenzeitlich liegt eine Abstimmungsfassung des Glossars mit knapp 60 Begriffen vor, deren Definitionen sich teils deutlich von denen des Eval-Wiki: Glossar der Evaluation unterscheiden (bspw. bei verschiedenen Resultaten eines Programms). Eine Verabschiedung nach einer erneuten Bearbeitung auf Grundlage von Kommentaren durch DeGEval-Mitglieder ist für die Mitgliederversammlung 2016 geplant.

– wie es im Vorwort zur zweiten Printversion (Univation 2009, S. 8) heißt – „als [einen] Wettbewerber auf einem Markt um Klärung in der Kommunikation in und über Evaluation“. Wettbewerber deshalb, weil unser Glossar – neben Erläuterungen in der Evaluationsfachliteratur – im deutschsprachigen Raum nicht das einzige Angebot in dieser Hinsicht ist (jedoch das umfangreichste). Es gibt verschiedene kleinere Glossare, die sich zumeist auf die Evaluation in ausgewählten Gegenstandsfeldern beziehen, wie bspw. das Glossar zu Evaluationsbegriffen des Schweizerischen Bundesamtes für Gesundheit (2005)¹² oder das Glossar entwicklungspolitischer Schlüsselbegriffe aus den Bereichen Evaluierung und ergebnisorientiertes Management der OECD (2009)¹³, wobei es sich bei Letzterem um eine Übersetzung des englischsprachigen Originals handelt. Eine umfangreichere Quelle stellt das ‚Online-Wörterbuch Evaluation‘ von Jan Hense dar, der darin auch einige zentrale Evaluatoreninnen/Evaluatoren mit ihren Beiträgen zur Theorie und Methodologie der Evaluation vorstellt und häufig deren – untereinander teils inkompatibel – Begriffsprägungen gewissermaßen werktreu und damit historisierend nutzt.¹⁴

5. Ausblick und Einladung zur Mitarbeit

Das Eval-Wiki: Glossar der Evaluation wird durch uns kontinuierlich gepflegt und immer wieder um neue Begriffe aus der Eval-Wiki: Glossar Werkstatt erweitert. Geplant ist zudem, das Eval-Wiki insgesamt auszubauen und weitere Werkzeuge und Ressourcen für Evaluierende und Evaluationsinteressierte zu ergänzen (bspw. Checklisten zur Evaluation, Informationen zur Selbstevaluation). Wir weisen in unseren Produkten und im Rahmen unserer Arbeit immer wieder auf das Eval-Wiki hin und bemühen uns um eine hohe Sichtbarkeit der Ressource, um dem Anspruch, die Profession Evaluation weiterzuentwickeln, gerecht zu werden.

Wir möchten Sie als Evaluierende bzw. Evaluationsinteressierte ganz ausdrücklich ermuntern und einladen, sich an der Terminologiearbeit zu beteiligen! Schlagen Sie bspw. Begriffe vor, die definiert werden sollten, oder kommentieren Sie bestehende Definitionen. Sie können dazu die Möglichkeiten im Eval-Wiki: Glossar der Evaluation oder in der Glossar-Werkstatt nutzen (Informationen zu Beteiligungs-

12 Downloadmöglichkeit unter <http://www.bag.admin.ch/evaluation/02357/02603/index.html?lang=de> [Stand: 19.02.2015]; hier wird auch auf das Eval-Wiki: Glossar der Evaluation als weitere Quelle verwiesen.

13 Siehe <http://www.oecd.org/development/evaluation/dcdndep/43184177.pdf> [Stand: 19.02.2015]

14 Siehe <http://www.evaluation.de/glossary> [Stand: 23.02.2015]. Einen Überblick über weitere Evaluationsglossare aus unterschiedlichen Sprachräumen gibt die Linksammlung von Lars Balzer unter <http://www.evaluation.lars-balzer.name/links/glossaries-dictionaries/> [Stand: 19.02.2015]. Englischsprachige Glossare, teils aus Fachpublikationen, sind aufgelistet in der Einleitung zum Printglossar 2009 (PDF unter http://www.univation.org/download/Univation_glossar_leseprobe.pdf [Stand: 19.02.2015]).

möglichkeiten finden Sie in den Portalen) oder auch die Diskussionsliste ‚forum evaluation‘, die durch Univation betreut wird.¹⁵ Wir freuen uns über Ihre Beiträge!

Literatur

- Balzer, Lars/Beywl, Wolfgang (2015): evaluiert. Planungsbuch für Evaluationen im Bildungsbereich. Bern: hep verlag.
- Beywl, Wolfgang (1988): Zur Weiterentwicklung der Evaluationsmethodologie. Grundlagen, Konzeption und Anwendung eines Modells der responsiven Evaluation. Frankfurt/M. u. a.: Verlag Peter Lang.
- Beywl, Wolfgang/Speer, Sandra/Kehr, Jochen (2004): Wirkungsorientierte Evaluation im Rahmen der Armut- und Reichtumsberichterstattung. Perspektivstudie. Im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung (BMGS). Köln. PDF-Download des Berichts: http://www.univation.org/download/Evaluation_der_Armut-_und_Reichtumsberichterstattung.pdf [Stand: 18.02.2015]
- Clinton, Janet (2014): The True Impact of Evaluation: Motivation for ECB. In: American Journal of Evaluation, Jg. 35, 1. S. 120–127.
- DeGEval (2002): Standards für Evaluation. Köln.
- Farrokhzad, Schahrzad/Mäder, Susanne (2014): Nutzenorientierte Evaluation. Ein Leitfaden für die Arbeitsfelder Integration, Vielfalt und Toleranz. Münster: Waxmann.
- Hattie, John A. C. (2013): Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von Visible Learning. Übersetzt und überarbeitet von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider Verlag.
- Hattie, John A. C./Clinton, Janet (2012): Physical activity is not related to performance at school. In: Arch Pediatr Adolesc Med, Jg. 166(7):678–9; author reply 679.
- Joint Committee on Standards for Educational Evaluation (JCSEE)/Sanders, James R. (Hrsg.) (2000): Handbuch der Evaluationsstandards. 2. Aufl. Opladen: Leske + Budrich.
- Patton, Michael Q. (1997): Utilization-focused evaluation. The new century text. 3rd ed. Thousand Oaks: Sage.
- Scriven, Michael (1991): Evaluation Thesaurus. Newbury Park: Sage.
- Scriven, Michael (2007): The Key Evaluation checklist. Western Michigan University. http://www.wmich.edu/evalctr/archive_checklists/kec_feb07.pdf [Stand: 04.03.2015]; in Deutsche übersetzt und kommentiert von Wolfgang Beywl und Daniela Schroeter unter http://www.univation.org/download/Key_Evaluation_Checklist_ger.pdf [Stand: 04.03.2015]
- Strahm, Peter (2008): Qualität durch systematisches Feedback: Grundlagen, Einblicke und Werkzeuge. Impulse zur Schulentwicklung. Bern: Schulverlag blmv.
- Thom, Norbert/Ritz, Adrian (2008): Public Management,. Innovative Konzepte zur Führung im öffentlichen Sektor. 4., aktualisierte Auflage. Wiesbaden: Gabler.
- Univation (Hrsg.) (2004): Das A-B-C der wirkungsorientierten Evaluation. Glossar – Deutsch/Englisch – der wirkungsorientierten Evaluation. Köln.
- Univation (Hrsg.) (2009): Das A-B-C der wirkungsorientierten Evaluation. Glossar – Deutsch/Englisch – der wirkungsorientierten Evaluation. 2., vollständig bearbeitete und ergänzte Auflage. Köln.

¹⁵ Informationen zum ‚forum evaluation‘ finden sich unter <http://www.univation.org/forum-evaluation> [Stand: 23.02.2015]

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

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